METHOD FOR THE GENERATION OF COMPACT TALE-NUCLEASES AND USES THEREOF

Field of the invention

The present invention relates to a method for the generation of compact Transcription Activator-Like Effector Nucleases (TALENs) that can efficiently target and process double-stranded DNA. More specifically, the present invention concerns a method for the creation of TALENs that consist of a single TALE DNA binding domain fused to at least one catalytic domain such that the active entity is composed of a single polypeptide chain for simple and efficient vectorization and does not require dimerization to target a specific single double-stranded DNA target sequence of interest and process DNA nearby said DNA target sequence. The present invention also relates to compact TALENs, vectors, compositions and kits used to implement the method.

Background of the invention

Mammalian genomes constantly suffer from various types of damage, of which double-strand breaks (DSBs) are considered the most dangerous (Haber 2000). Repair of DSBs can occur through diverse mechanisms that can depend on cellular context. Repair via homologous recombination (HR)is able to restore the original sequence at the break. Because of its strict dependence on extensive sequence homology, this mechanism is suggested to be active mainly during the S and G2 phases of the cell cycle where the sister chromatids are in close proximity (Sonoda, Hochegger et al. 2006). Single-strand annealing (SSA) is another homology-dependent process that can repair DSBs between direct repeats and thereby promotes deletions (Paques and Haber 1999). Finally, non-homologous end joining (NHEJ) of DNA is a major pathway for the repair of DSBs that can function throughout the cell cycle and does not depend on homologous recombination (Moore and Haber 1996; Haber 2008). NHEJ seems to comprise at least two different components: (i) a pathway that consists mostly in the direct re-joining of DSB ends, and which depends on the XRCC4, Lig4 and Ku proteins, and; (ii) an alternative NHEJ pathway, which does not depend on XRCC4, Lig4 and Ku, and is especially error-prone, resulting mostly in deletions, with the junctions occurring between micro-homologies (Frank, Sekiguchi et al. 1998; Gao, Sun et al. 1998; Guirouilh-Barbat, Huck et al. 2004; Guirouilh-Barbat, Rass et al. 2007; Haber 2008; McVey and Lee 2008).

Homologous gene targeting (HGT), first described over 25 years ago (Hinnen, Hicks et al. 1978; Orr-Weaver, Szostak et al. 1981; Orr-Weaver, Szostak et al. 1983; Rothstein 1983), was one of the first methods for rational genome engineering and remains to this day a standard for the generation of

engineered cells or knock-out mice (Capecchi 2001). An inherently low efficiency has nevertheless prevented it from being used as a routine protocol in most cell types and organisms. To address these issues, an extensive assortment of rational approaches has been proposed with the intent of achieving greater than 1% targeted modifications. Many groups have focused on enhancing the efficacy of HGT, with two major disciplines having become apparent: (i) so-called "matrix optimization" methods, essentially consisting of modifying the targeting vector structure to achieve maximal efficacy, and; (ii) methods involving additional effectors to stimulate HR, generally sequence-specific endonucleases. The field of matrix optimization has covered a wide range of techniques, with varying degrees of success (Russell and Hirata 1998; Inoue, Dong et al. 2001; Hirata, Chamberlain et al. 2002; Taubes 2002; Gruenert, Bruscia et al. 2003; Sangiuolo, Scaldaferri et al. 2008; Bedayat, Abdolmohamadi et al. 2010). Stimulation of HR via nucleases, on the other hand, has repeatedly proven efficient (Paques and Duchateau 2007; Carroll 2008).

For DSBs induced by biological reagents, e.g. meganucleases, ZFNs and TALENs (see below), which cleave DNA by hydrolysis of two phosphodiester bonds, the DNA can be rejoined in a seamless manner by simple re-ligation of the cohesive ends. Alternatively, deleterious insertions or deletions (indels) of various sizes can occur at the breaks, eventually resulting in gene inactivation (Liang, Han et al. 1998; Lloyd, Plaisier et al. 2005; Doyon, McCammon et al. 2008; Perez, Wang et al. 2008; Santiago, Chan et al. 2008; Kim, Lee et al. 2009; Yang, Djukanovic et al. 2009). The nature of this process, which does not rely on site-specific or homologous recombination, gives rise to a third targeted approach based on endonuclease-induced mutagenesis. This approach, as well as the related applications, may be simpler than those based on homologous recombination in that (a) one does not need to introduce a repair matrix, and; (b) efficacy will be less cell-type dependant (in contrast to HR, NHEJ is probably active throughout the cell cycle (Delacote and Lopez 2008). Targeted mutagenesis based on NEHJ has been used to trigger inactivation of single or even multiple genes in immortalized cell lines (Cost, Freyvert et al. 2010; Liu, Chan et al. 2010). In addition, this method opens new perspectives for organisms in which the classical HR-based gene knock-out methods have proven inefficient, or at least difficult to establish (Doyon, McCammon et al. 2008; Geurts, Cost et al. 2009; Shukla, Doyon et al. 2009; Yang, Djukanovic et al. 2009; Gao, Smith et al. 2010; Mashimo, Takizawa et al. 2010; Menoret, Iscache et al. 2010).

Over the last 15 years, the use of meganucleases to successfully induce gene targeting has been well documented, starting from straightforward experiments involving wild-type I-SceI to more refined work involving completely re-engineered enzymes (Stoddard, Scharenberg et al. 2007; Galetto,

Duchateau et al. 2009; Marcaida, Munoz et al. 2010; Arnould, Delenda et al. 2011). Meganucleases, also called homing endonucleases (HEs), can be divided into five families based on sequence and structure motifs: LAGLIDADG, GIY-YIG, HNH, His-Cys box and PD-(D/E)XK (Stoddard 2005; Zhao, Bonocora et al. 2007). Structural data are available for at least one member of each family. The most well studied family is that of the LAGLIDADG proteins, with a considerable body of biochemical, genetic and structural work having established that these endonucleases could be used as molecular tools (Stoddard, Scharenberg et al. 2007; Arnould, Delenda et al. 2011). Member proteins are composed of domains that adopt a similar $\alpha\beta\beta\alpha\beta\beta\alpha$ fold, with the LAGLIDADG motif comprising the terminal region of the first helix and not only contributing to a bipartite catalytic center but also forming the core subunit/subunit interaction (Stoddard 2005). Two such α/β domains assemble to form the functional protein, with the β-strands in each creating a saddle- shaped DNA binding region. The spatial separation of the catalytic center with regions directly interacting with the DNA has allowed for specificity re-engineering (Seligman, Chisholm et al. 2002; Sussman, Chadsey et al. 2004; Arnould, Chames et al. 2006; Doyon, Pattanayak et al. 2006; Rosen, Morrison et al. 2006; Smith, Grizot et al. 2006; Arnould, Perez et al. 2007). In addition, whereas all known LAGLIDADG proteins analyzed to date act as "cleavases" to cut both strands of the target DNA, recent progress has been made in generating "mega-nickases" that cleave only one strand (Niu, Tenney et al. 2008; McConnell Smith, Takeuchi et al. 2009). Such enzymes can in principle provide similar levels of targeted induced HR with a minimization in the frequency of NHEJ.

Although numerous engineering efforts have focused on LAGLIDADG HEs, members from two other families, GIY-YIG and HNH, are of particular interest. Biochemical and structural studies have established that in both families, member proteins can adopt a bipartite fold with distinct functional domains: (1) a catalytic domain responsible mainly for DNA cleavage, and; (2) a DNA-binding domain to provide target specificity (Stoddard 2005; Marcaida, Munoz et al. 2010). The related GIY-YIG HEs I-TevI and I-BmoI have been exploited to demonstrate the interchangeability of the DNA-binding region for these enzymes (Liu, Derbyshire et al. 2006). Analysis of the I-BasI HE revealed that although the N-terminal catalytic domain belongs to the HNH family, the C-terminal DNA-binding region resembles the intron-encoded endonuclease repeat motif (IENR1) found in endonucleases of the GIY-YIG family (Landthaler and Shub 2003). The catalytic head of I-BasI has sequence similarity to those of the HNH HEs I-HmuI, I-HmuII and I-TwoI, all of which function as strand-specific nickases (Landthaler, Begley et al. 2002; Landthaler and Shub 2003; Landthaler, Lau et al. 2004; Shen, Landthaler et al. 2004; Landthaler, Shen et al. 2006).

Whereas the above families of proteins contain sequence-specific nucleases, the HNH motif has also been identified in nonspecific nucleases such the *E.coli* colicins (e.g. ColE9 and ColE7), EndA from *S.pneumoniae*, NucA from *Anabaena* and CAD (Midon, Schafer et al. 2011). As well as having the HNH motif, several of these nucleases contain the signature DRGH motif and share structural homology with core elements forming the ββα–Me-finger active site motif. Mutational studies of residues in the HNH/DRGH motifs have confirmed their role in nucleic acid cleavage activity (Ku, Liu et al. 2002; Doudeva, Huang et al. 2006; Eastberg, Eklund et al. 2007; Huang and Yuan 2007). Furthermore, the DNA binding affinity and sequence preference for ColE7 could be effectively altered (Wang, Wright et al. 2009). Such detailed studies illustrate the potential in re-engineering nonspecific nucleases for targeted purposes.

Zinc-finger nucleases (ZFNs), generated by fusing Zinc-finger-based DNA-binding domains to an independent catalytic domain via a flexible linker (Kim, Cha et al. 1996; Smith, Berg et al. 1999; Smith, Bibikova et al. 2000), represent another type of engineered nuclease commonly used to stimulate gene targeting. The archetypal ZFNs are based on the catalytic domain of the Type IIS restriction enzyme FokI and have been successfully used to induce gene correction, gene insertion, and gene deletion. Zinc Finger-based DNA binding domains are made of strings of 3 or 4 individual Zinc Fingers, each recognizing a DNA triplet (Pabo, Peisach et al. 2001). In theory, one of the major advantages of ZFNs is that they are easy to design, using combinatorial assembly of preexisting Zinc Fingers with known recognition patterns (Choo and Klug 1994; Choo and Klug 1994; Kim, Lee et al. 2009). However, close examination of high resolution structures shows that there are actually cross-talks between units (Elrod-Erickson, Rould et al. 1996), and several methods have been used to assemble ZF proteins by choosing individual Zinc Fingers in a context dependant manner (Greisman and Pabo 1997; Isalan and Choo 2001; Maeder, Thibodeau-Beganny et al. 2008; Ramirez, Foley et al. 2008) to achieve better success rates and reagents of better quality.

Recently, a new class of chimeric nuclease using a FokI catalytic domain has been described (Christian, Cermak et al. 2010; Li, Huang et al. 2011). The DNA binding domain of these nucleases is derived from Transcription Activator Like Effectors (TALE), a family of proteins used in the infection process by plant pathogens of the Xanthomonas genus. In these DNA binding domains, sequence specificity is driven by a series of 33-35 amino acids repeats, differing essentially by two positions (Boch, Scholze et al. 2009; Moscou and Bogdanove 2009). Each base pair in the DNA target is contacted by a single repeat, with the specificity resulting from the two variant amino acids of the repeat (the so-called repeat variable dipeptide, RVD). The apparent modularity of these DNA binding

domains has been confirmed to a certain extent by modular assembly of designed TALE-derived protein with new specificities (Boch, Scholze et al. 2009; Moscou and Bogdanove 2009). However, one cannot yet rule out a certain level of context dependence of individual repeat/base recognition patterns, as was observed for Zinc Finger proteins (see above). Furthermore, it has been shown that natural TAL effectors can dimerize (Gurlebeck, Szurek et al. 2005) and how this would affect a "dimerization-based" TALE-derived nuclease is currently unknown.

The functional layout of a FokI-based TALE-nuclease (TALEN) is essentially that of a ZFN, with the Zinc-finger DNA binding domain being replaced by the TALE domain (Christian, Cermak et al. 2010; Li, Huang et al. 2011). As such, DNA cleavage by a TALEN requires two DNA recognition regions flanking an unspecific central region. This central "spacer" DNA region is essential to promote catalysis by the dimerizing FokI catalytic domain, and extensive effort has been placed into optimizing the distance between the DNA binding sites (Christian, Cermak et al. 2010; Miller, Tan et al. 2011). The length of the spacer has been varied from 14 to 30 base pairs, with efficiency in DNA cleavage being interdependent with spacer length as well as TALE scaffold construction (i.e. the nature of the fusion construct used). It is still unknown whether differences in the repeat region (i.e. RVD type and number used) have an impact on the DNA "spacer" requirements or on the efficiency of DNA cleavage by TALENs. Nevertheless, TALE-nucleases have been shown to be active to various extents in cell-based assays in yeast, mammalian cells and plants (Christian, Cermak et al. 2010; Li, Huang et al. 2011; Mahfouz, Li et al. 2011; Miller, Tan et al. 2011).

The inventors have developed a new type of TALEN that can be engineered to specifically recognize and process target DNA efficiently. These novel "compact TALENs" (cTALENs) do not require dimerization for DNA processing activity, thereby alleviating the need for "dual" target sites with intervening DNA "spacers". Furthermore, the invention allows for generating several distinct types of enzymes that can enhance separate DNA repair pathways (HR vs. NHEJ).

Brief summary of the invention

The present invention relates to a method to generate compact Transcription Activator-Like Effector Nucleases (TALENs) composed of a single polypeptide chain that do not require dimerization to target a specific single double-stranded DNA target sequence of interest and process DNA nearby said single double-stranded DNA target sequence of interest. The present invention also concerns the creation of functional single polypeptide fusion proteins for simple and efficient vectorization. In

another aspect, the present invention relates to compact TALENs comprising at least an enhancer domain wherein said enhancer domain enhances the DNA processing efficiency of said compact TALENS nearby a single double-stranded DNA target sequence of interest. The present invention also relates to compact TALENS, vectors, compositions and kits used to implement the method. The present invention also relates to methods for use of said compact TALENs according to the invention for various applications ranging from targeted DNA cleavage to targeted gene regulation. The methods according to the present invention can be used in various fields ranging from the creation of transgenic organisms to treatment of genetic diseases.

Brief description of the figures

In addition to the preceding features, the invention further comprises other features which will emerge from the description which follows, as well as to the appended drawings. A more complete appreciation of the invention and many of the attendant advantages thereof will be readily obtained as the same becomes better understood by reference to the following Figures in conjunction with the detailed description below.

Figure 1: Endonuclease-induced gene targeting approaches. Upon cleavage, DNA repair mechanisms may result in one of several outcomes. (A) When a double-strand break is targeted between two direct repeats, HR can result in the deletion of one repeat together with the intervening sequence. Gene insertion (B) or correction (C) can be achieved by the introduction of a DNA repair matrix containing sequences homologous to the endogenous sequence surrounding the DNA break. Mutations can be corrected either at or distal to the break, with the frequency of correction decreasing with increasing distance. (D) The misrepair of DNA ends by error-prone NHEJ can result in insertions or deletions of various sizes, leading to gene inactivation.

Figure 2: Sequences of target DNA recognized by I-CreI and I-TevI. C1234 (SEQ ID NO: 3) represents the partially symmetric natural DNA sequence recognized and cleaved by the wild-type I-CreI meganuclease. C1221 (SEQ ID NO: 2) represents an artificial palindromic DNA sequence, derived from C1234 (SEQ ID NO: 3), also recognized and cleaved by I-CreI (SEQ ID NO: 1). Nucleotides are numbered outward (-/+) from the center of the target. DNA cleavage occurs on either side of the underlined sequence to generate 4-nucleotide 3' overhanging ends. For I-CreI-based meganucleases, the nature of the nucleotides at positions -2 to +2 can potentially interfere with the cleavage activity of the protein. Tev (SEQ ID NO: 105) represents the asymmetric DNA sequence recognized and

cleaved by the wild-type I-TevI meganuclease. Nucleotide numbering is relative to the intron-insertion site of the natural target sequence. Cleavage by I-TevI occurs on either side of the underlined sequence to generate 2-nucleotide 3' overhanging ends.

Figure 3: Sequences of the target DNAs recognized by TALEN and compact TALEN constructs. Target DNAs for the engineered compact TALENS cTN-Avr and cTN-Pth are based on the naturally occurring asymmetric sequences AvrBs3 (19 bp) [in bT1-Avr (SEQ ID NO: 136) and bT2-Avr (SEQ ID NO: 137) baseline protein scaffolds] and PthXo1 (25 bp) [in bT1-Pth (SEQ ID NO: 138) and bT2-Pth (SEQ ID NO: 139) baseline protein scaffolds], respectively. For each sequence, nucleotides are numbered outward (-/+) from the anchoring T (position -1). Sequences shown are directly contacted by the protein to provide target specificity. Wild-type Repeat Variable Dipeptides (RVDs) correspond to the dipeptides found in the repeats of the naturally occurring effector proteins targeted to each sequence. Cipher RVDs are based on the subset of dipeptide/nucleotide pairs listed. Artificial RVDs are derived by direct readout of the underlying DNA sequence using the cipher RVD code (SEQ ID NO: 245 to 249).

Figure 4: Schematic of meganuclease fusion configurations. Fusion constructs are optimized to address or overcome distinct problems. (A) The addition of two catalytic domains to an active meganuclease can not only enhance cleavage activity (e.g. three chances to effect DNA cleavage per binding event) but can also promote sequence alterations by error-prone NHEJ since small sections of DNA are excised for each pair of cleavage events. (B) When specificity reengineering precludes maintaining cleavage activity of the meganuclease, the attached catalytic domains provide the necessary strand cleavage function. (C) and (D) represent instances of (A) and (B), respectively, when only one catalytic domain is tolerated per fusion protein (e.g. either as an N- or C-terminal fusion or in the context of a single-chain molecule). In all cases, the catalytic domain envisioned can be either a cleavase (ability to cleave both strands of the DNA) or a nickase (cleavage of only a single DNA strand) depending on the application. Fusion junctions (N- vs. C-terminal) and linker designs can vary with the application. Components of the fusion proteins are listed in the legend.

Figure 5: Schematic of cTALEN configurations. Compact TALENs are designed to alleviate the need for multiple independent protein moieties when targeting a DNA cleavage event. Importantly, the requisite "spacer" region and dual target sites essential for the function of current classical TALENs are unnecessary. In addition, since the catalytic domain does not require specific DNA contacts,

there are no restrictions on regions surrounding the core TALE DNA binding domain. (A) N-terminal fusion construct to promote HR via a standard (cleavase domain) or conservative (nickase domain) repair pathway. (B) C-terminal fusion construct with properties as in (A). (C) The attachment of two catalytic domains to both ends of the TALE allows for dual cleavage with enhancement in NHEJ. Fusion junctions (N- vs. C-terminal) and linker designs can vary with the application. Components of the fusion proteins are listed in the legend.

Figure 6: Schematic of enhanced cTALEN configurations. Compact TALENs can be enhanced through the addition of a domain to promote existing or alternate activities. As each end of the TALE DNA binding domain is amenable to fusion, the order (N- v.s C-terminal) of addition of the catalytic and enhancer domains can vary with the application. (A) A standard cTALEN with a C-terminal enhancer domain. (B) The enhancer domain is fused to the cTALEN via the N-terminus of the catalytic domain. Such a configuration can be used to assist and/or anchor the catalytic domain near the DNA to increase cleavage activity. (C) The enhancer domain is sandwiched between the catalytic domain and TALE DNA binding domain. The enhancer domain can promote communication between the flanking domains (i.e. to assist in catalysis and/or DNA binding) or can be used to overcome the requisite T nucleotide at position -1 of all TALE-based targets. (D) The enhancer domain is used to functionally replace the natural TALE protein N-terminal region. (E) The enhancer domain is used to functionally replace the natural TALE protein C-terminal region. Fusion junctions (N- vs. C-terminal) and linker designs can vary with the application. Components of the fusion proteins are listed in the legend.

Figure 7: Schematic of trans cTALEN configurations. Compact TALENs can be combined with auxiliary enhancer domains to promote alternate activities. Auxiliary domains provide an additional function that is not essential to the cTALEN activity. (A) A standard cTALEN with an N-terminal nickase catalytic domain becomes a "cleavase" via the separate addition of the auxiliary domain. (B) In some instances, the need to target the specificity of the auxiliary domain is necessary. Such a configuration can achieved via a TALE fusion and can be used to assist and/or anchor the auxiliary domain near the DNA to increase activity of the cTALEN. (C) The targeted auxiliary domain is provided either before or after the cTALEN to perform an independent task. Communication between the fusion proteins is not necessary. Fusion junctions (N- vs. C-terminal) and linker designs can vary with the application. Components of the fusion proteins are listed in the legend.

Figure 8 : Schematic of DNA cleavage, in vivo re-ligation and other repair pathways. In cells, cleavage by peptidic rare-cutting endonucleases usually results in a DNA double strand break (DSB) with cohesive ends. For example, meganucleases from the LAGLIDADG family, such as I-SceI and I-

CreI, produce DSBs with 3' overhangs. These cohesive ends can be re-ligated *in vivo* by NHEJ, resulting in seamless repair, and in the restoration of a cleavable target sequence, which can in turn be processed again by the same endonuclease. Thus, a series of futile cycles of cleavage and re-ligation events can take place. Imprecise NHEJ or homologous recombination can alter or remove the cleavage site, resulting in cycle exit; this can also apply to compact TALENs and enhanced compact TALENs according to the present invention (**A**). Two other ways can also stop the process: (i) Chromosome loss can occur as the consequence of failure to repair the DSB; (ii) a loss of nuclease (degradation, dilution, cell division, etc...). **B-E**: Consequences of cleavage of additional phosphodiester bonds. The addition of a single nickase activity (**B**) or of two nickase activities affecting the same strand (**C**) would result in a single strand gap, and suppress the cohesive ends, which could in turn affect the spectrum of events. Addition of two nickase activities affecting opposite strands (**D**) or of a new cleavase activity generating a second DSB (**E**) would result in a double strand gap; as a consequence, perfect re-ligation is no longer possible, and one or several alternative repair outcomes could be stimulated. The current figure makes no assumption regarding the relative frequencies of these alternative outcomes (imprecise NHEJ, homologous recombination, others...). Solid triangles represent hydrolysis of phosphodiester bonds.

Figure 9: Activity of TALE-AvrBs3::TevI in yeast (37°C). The negative control consists in a TALEN without any RVDs. n.d. indicates no detectable activity, + indicates an activity over 0.3 in yeast assay and +++ indicates an activity over 0.7 in yeast assay (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

Figure 10: Activity of TALE-AvrBs3::TevI in mammalian cells. (Extrachromosomic assay in CHO-K1). pCLS8993 (SEQ ID NO: 194) is represented by a black bar and pCLS8994 (SEQ ID NO: 195) is represented by a dark grey bar. Negative control (empty vector) by a white bar and positive control (I-SceI meganuclease) by a light grey bar. Data are normalized relative to the positive control.

Figure 11: Activity of TALE-AvrBs3::NucA in yeast (37°C). The negative control is a target lacking a recognition site (neg. ctrl. : SEQ ID NO: 228). Compact is a target having only one recognition site (SEQ ID NO: 224). n.d. indicates no detectable activity, + indicates an activity over 0.3 in yeast assay at 37°C; ++, activity over 0.5 in yeast assay at 37°C and +++ activity over 0.7 in yeast assay at 37°C (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

Figure 12: Activity of TALE-AvrBs3::ColE7 in yeast (37°C). The negative control is a target lacking a recognition site (neg. ctrl. : SEQ ID NO: 228). Compact is a target having only one recognition site (SEQ ID NO: 224). n.d. indicates no detectable activity, + indicates an activity over 0.3 in yeast assay at 37°C; ++, activity over 0.5 in yeast assay at 37°C and +++ activity over 0.7 in yeast assay at 37°C (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

Figure 13: Schematic and details of a TALE::CreI prototypical compact TALEN. This class of compact TALEN targets a bipartite recognition sequence comprised of the TALE DNA binding site proximal to a meganuclease target site. The engineered TCRB02-A meganuclease site is shown along with details of the RVDs and DNA sequences recognized by the TALE moiety. A region of the T cell receptor B gene is presented, highlighting the endogenous layout of the TALE::CreI-based compact TALEN hybrid target site.

Figure 14: Activity of TALE::scTB2aD01-based constructs in yeast (30°C). The layouts of the various hybrid targets are shown, starting (5') with the region recognized by the TALE DNA binding domain in uppercase, the unspecific spacer region in lowercase and the meganuclease target site in underlined uppercase characters. Activity in yeast is illustrated for select representative constructs. n.d. indicates no detectable activity, + indicates an activity over 0.3 in yeast assay at 30°C; ++, activity over 0.5 in yeast assay at 30°C and +++ activity over 0.7 in yeast assay at 30°C (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

Figure 15: Western blot of TALE::scTB2aD01-based constructs. Constructs were expressed in HEK293 cells and total protein extracts were prepared 48 hours post-transfection. Protein was detected using a polyclonal anti-I-CreI antibody.

Figure 16: Toxicity of TALE::scTB2aD01-based constructs in CHOK1 cells. Cytotoxicty is based on detectable levels of GFP expressed in living cells, on day 1 vs day 6, relative to a standard control (transfection of empty plasmid).

Figure 17: NHEJ activity of TALE::scTB2aD01-based constructs in HEK293 cells. A post-transfection PCR-based analysis of genomic DNA is used to assess activity *in vivo*. Cleavage of mismatched DNA sequences by T7 endonuclease is indicative of NHEJ events resulting from the activity of the cTALEN or meganuclease at the targeted locus.

Figure 18: Activity of TevI::TALE-AvrBs3 +/- TALE-RagT2-R::TevI in yeast (37°C). The negative control is a target lacking a recognition site (neg. ctrl. : SEQ ID NO: 228). n.d. indicates no detectable activity, + indicates an activity over 0.3 in yeast assay at 37°C; ++, activity over 0.5 in yeast assay at 37°C and +++ activity over 0.7 in yeast assay at 37°C (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

Figure 19: Activity of TALE::SnaseSTAAU in yeast (37°C). The negative control is a target lacking recognition sites. Compact is a target having only one recognition site (SEQ ID NO: 224). n.d. indicates no detectable activity at 37°C, +/- indicated an activity above 0.3 in yeast assay at 37°C; + indicated an activity over 0.3 in yeast assay at 37°C; ++ indicated an activity over 0.5 in yeast assay at 37°C; +++ indicated an activity over 0.75 in yeast assay at 37°C (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

Figure 20: Activity of TALE::ColE7 with various polypeptide linker in yeast (37°C). Compact is a target having only one recognition site (SEQ ID NO: 224). n.d. indicates no detectable activity at 37°C, + indicated an activity over 0.3 in yeast assay at 37°C; ++ indicated an activity over 0.5 in yeast assay at 37°C; +++ indicated an activity over 0.75 in yeast assay at 37°C (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

Table 6: List of AvrBs3 targets with various spacer lengths (SEQ ID NO: 157 to 192).

Table 7: List of AvrBs3 targets with various spacer lengths (SEQ ID NO: 157 to 192) including a target with only one recognition site (compact, SEQ ID NO: 224) and a negative control target (neg. ctrl., SEQ ID NO: 228) consisting in a target without any recognition site.

Table 13: List of hybrid RagT2-R/AvrBs3 targets with various spacer lengths (SEQ ID NO: 315 to 350).

Detailed description of the invention

Unless specifically defined herein, all technical and scientific terms used have the same meaning as commonly understood by a skilled artisan in the fields of gene therapy, biochemistry, genetics, and molecular biology.

All methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, with suitable methods and materials being described herein. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will prevail. Further, the materials, methods, and examples are illustrative only and are not intended to be limiting, unless otherwise specified.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See, for example, Current Protocols in Molecular Biology (Frederick M. AUSUBEL, 2000, Wiley and son Inc, Library of Congress, USA); Molecular Cloning: A Laboratory Manual, Third Edition, (Sambrook et al, 2001, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press); Oligonucleotide Synthesis (M. J. Gait ed., 1984); Mullis et al. U.S. Pat. No. 4,683,195; Nucleic Acid Hybridization (B. D. Harries & S. J. Higgins eds. 1984); Transcription And Translation (B. D. Hames & S. J. Higgins eds. 1984); Culture Of Animal Cells (R. I. Freshney, Alan R. Liss, Inc., 1987); Immobilized Cells And Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide To Molecular Cloning (1984); the series, Methods In ENZYMOLOGY (J. Abelson and M. Simon, eds.-in-chief, Academic Press, Inc., New York), specifically, Vols.154 and 155 (Wu et al. eds.) and Vol. 185, "Gene Expression Technology" (D. Goeddel, ed.); Gene Transfer Vectors For Mammalian Cells (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); Immunochemical Methods In Cell And Molecular Biology (Mayer and Walker, eds., Academic Press, London, 1987); Handbook Of Experimental Immunology, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds., 1986); and Manipulating the Mouse Embryo, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

In a first aspect, the present invention relates to a method to generate compact Transcription Activator-Like Effector Nucleases (cTALENs) composed of a single polypeptide chain that do not require dimerization to target a specific single double-stranded DNA target sequence of interest and process DNA nearby said single double-stranded DNA target sequence of interest.

According to a first aspect of the present invention is a method to generate compact Transcription Activator-Like Effector Nucleases (cTALENs) comprising the steps of:

(i) Engineering a core TALE scaffold (a) comprising different sets of Repeat Variable Dipeptide regions (RVDs) to change DNA binding specificity and target a specific single double-stranded DNA target sequence of interest, (b) onto which a selection of catalytic domains can be attached to effect DNA processing;

(ii) Determining or engineering at least one catalytic domain wherein said catalytic domain is capable of processing DNA nearby said single double-stranded DNA target sequence of interest when fused to said engineered core TALE scaffold from (i);

(iii) Optionally determining or engineering a peptidic linker to fuse said catalytic domain from (ii) to said engineered core TALE scaffold from (i);

thereby obtaining a compact TALEN entity composed of a single polypeptide chain that does not require dimerization to target a specific single double-stranded DNA target sequence of interest and process DNA nearby said single double-stranded DNA target sequence of interest. In other words, the compact TALEN according to the present invention is an active entity unit able, by itself, to target only one specific single double-stranded DNA target sequence of interest through one DNA binding domain and to process DNA nearby said single double-stranded DNA target sequence of interest.

In another embodiment, is a method for targeting and processing a double-stranded DNA, comprising:

(a) Selecting one DNA target sequence of interest on one strand of a double-stranded DNA;

(b) Providing a unique compact TALEN monomer comprising:

(i) One core TALE scaffold comprising Repeat Variable Dipeptide regions (RVDs) having DNA binding specificity onto said DNA target sequence of interest;

(ii) At least one catalytic domain wherein said catalytic domain is capable of processing DNA a few base pairs away from said DNA target sequence of interest when fused to the C and/or N terminal of said core TALE scaffold from (i);

(iii) Optionally one peptidic linker to fuse said catalytic domain from (ii) to said core TALE scaffold from (i) when needed;

wherein said compact TALEN monomer is assembled to bind and process said double stranded DNA without requiring dimerization;

(c) Contacting said double-stranded DNA with said unique monomer such that the double-stranded is processed a few base pairs away in 3' and/or 5' direction(s) from said one strand target sequence.

In another embodiment, said engineered core TALE scaffold according to the present invention comprises an additional N-terminal domain resulting in an engineered core TALE scaffold sequentially comprising a N-terminal domain and different sets of Repeat Variable Dipeptide regions (RVDs) to change DNA binding specificity and target a specific single double-stranded DNA target sequence of interest, onto which a selection of catalytic domains can be attached to effect DNA processing.

In another embodiment, said engineered core TALE scaffold according to the present invention comprises an additional C-terminal domain resulting in an engineered core TALE scaffold sequentially comprising different sets of Repeat Variable Dipeptide regions (RVDs) to change DNA binding specificity and target a specific single double-stranded DNA target sequence of interest and a C-terminal domain, onto which a selection of catalytic domains can be attached to effect DNA processing.

In another embodiment, said engineered core TALE-scaffold according to the present invention comprises additional N-terminus and a C-terminal domains resulting in an engineered core TALE scaffold sequentially comprising a N-terminal domain, different sets of Repeat Variable Dipeptide regions (RVDs) to change DNA binding specificity and target a specific single double-stranded DNA target sequence of interest and a C-terminal domain, onto which a selection of catalytic domains can be attached to effect DNA processing. In another embodiment, said engineered core TALE-scaffold according to the present invention comprises the protein sequences selected from the group consisting of ST1 (SEQ ID NO: 134) and ST2 (SEQ ID NO: 135). In another embodiment, said engineered TALE-scaffold comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group consisting of SEQ ID NO: 134 and SEQ ID NO: 135.

In another embodiment, said engineered core TALE-scaffold according to the present invention comprises the protein sequences selected from the group consisting of bT1-Avr (SEQ ID NO: 136), bT2-Avr (SEQ ID NO: 137), bT1-Pth (SEQ ID NO: 138) and bT2-Pth (SEQ ID NO: 139). In another

embodiment, said engineered TALE-scaffold comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group consisting of SEQ ID NO: 136 to SEQ ID NO: 139.

In a preferred embodiment according to the method of the present invention, said additional N-terminus and C-terminal domains of engineered core TALE scaffold are derived from natural TALE. In a more preferred embodiment said additional N-terminus and C-terminal domains of engineered core TALE scaffold are derived from natural TALE like AvrBs3, PthXo1, AvrHah1, PthA, Tal1c as non-limiting examples. In another more preferred embodiment, said additional N-terminus and/or said C-terminal domains are truncated forms of respective N-terminus and/or said C-terminal domains of natural TALE like AvrBs3, PthXo1, AvrHah1, PthA, Tal1c as non-limiting examples from which they are derived. In a more preferred embodiment, said additional N-terminus and C-terminal domains sequences of engineered core TALE scaffold are selected from the group consisting of ST1 SEQ ID NO: 134 and ST2 SEQ ID NO: 135 as respectively exemplified in baseline protein scaffolds bT1-Avr (SEQ ID NO: 136) or bT1-Pth (SEQ ID NO: 138) and bT2-Avr (SEQ ID NO: 137) or bT2-Pth (SEQ ID NO: 139).

In another embodiment, each RVD of said core scaffold is made of 30 to 42 amino acids, more preferably 33 or 34 wherein two critical amino acids located at positions 12 and 13 mediates the recognition of one nucleotide of said nucleic acid target sequence; equivalent two critical amino acids can be located at positions other than 12 and 13 specialy in RVDs taller than 33 or 34 amino acids long. Preferably, RVDs associated with recognition of the different nucleotides are HD for recognizing C, NG for recognizing T, NI for recognizing A, NN for recognizing G or A, NS for recognizing A, C, G or T, HG for recognizing T, IG for recognizing T, NK for recognizing G, HA for recognizing C, ND for recognizing C, HI for recognizing C, HN for recognizing G, NA for recognizing G, SN for recognizing G or A and YG for recognizing T, TL for recognizing A, VT for recognizing A or G and SW for recognizing A. More preferably, RVDs associated with recognition of the nucleotides C, T, A, G/A and G respectively are selected from the group consisting of NN or NK for recognizing G, HD for recognizing C, NG for recognizing T and NI for recognizing A, TL for recognizing A, VT for recognizing A or G and SW for recognizing A. In another embodiment, RVDS associated with recognition of the nucleotide C are selected from the group consisting of N* and RVDS associated with recognition of the nucleotide T are selected from the group consisting of N* and H*, where * denotes a gap in the repeat sequence that corresponds to a lack of amino acid residue at the second position of the RVD. In another embodiment, critical amino acids 12 and 13 can be mutated towards other amino acid

residues in order to modulate their specificity towards nucleotides A, T, C and G and in particular to enhance this specificity. By other amino acid residues is intended any of the twenty natural amino acid residues or unnatural amino acids derivatives.

In another embodiment, said core scaffold of the present invention comprises between 8 and 30 RVDs. More preferably, said core scaffold of the present invention comprises between 8 and 20 RVDs; again more preferably 15 RVDs.

In another embodiment, said core scaffold comprises an additional single truncated RVD made of 20 amino acids located at the C-terminus of said set of RVDs, i.e. an additional C-terminal half-RVD. In this case, said core scaffold of the present invention comprises between 8.5 and 30.5 RVDs, ".5" referring to previously mentioned half-RVD (or terminal RVD, or half-repeat). More preferably, said core scaffold of the present invention comprises between 8.5 and 20.5 RVDs, again more preferably, 15.5 RVDs. In a preferred embodiment, said half-RVD is in a core scaffold context which allows a lack of specificity of said half-RVD toward nucleotides A, C, G, T. In a more preferred embodiment, said half-RVD is absent.

In another embodiment, said core scaffold of the present invention comprises RVDs of different origins. In a preferred embodiment, said core scaffold comprises RVDs originating from different naturally occurring TAL effectors. In another preferred embodiment, internal structure of some RVDs of the core scaffold of the present invention are constituted by structures or sequences originated from different naturally occurring TAL effectors. In another embodiment, said core scaffold of the present invention comprises RVDs-like domains. RVDs-like domains have a sequence different from naturally occurring RVDs but have the same function and / or global structure within said core scaffold of the present invention.

In another embodiment, said additional N-terminal domain of said engineered core TALE scaffold is an enhancer domain. In another embodiment, said enhancer domain is selected from the group consisting of Puf RNA binding protein or Ankyrin super-family, as non-limiting examples. In another embodiment, said enhancer domain sequence is selected from the group consisting of protein domains of SEQ ID NO: 4 and SEQ ID NO: 5, as non-limiting examples listed in Table 1, a functional mutant, a variant or a derivative thereof..

In another embodiment, said additional C-terminal domain of said engineered core TALE scaffold is an enhancer domain. In another embodiment, said enhancer domain is selected from the group

consisting of hydrolase/transferase of Pseudomonas Aeuriginosa family, the polymerase domain from the Mycobacterium tuberculosis Ligase D family, the initiation factor eIF2 from Pyrococcus family, the translation initiation factor Aif2 family, as non-limiting examples. In another embodiment, said enhancer domain sequence is selected from the group consisting of protein domains of SEQ ID NO: 6 to SEQ ID NO: 9, as non-limiting examples listed in Table 1, a functional mutant, a variant or a derivative thereof.

GENBANK/SWISS-PROT ID	NAME	SEQ ID NO	FASTA SEQUENCE
gi|262368139|pdb|3K5Q|	fem-3	4	>gi|262368139|pdb|3K5Q|A Chain A, Crystal Structure Of Fbf-2FBE COMPLEX SNNVLPTWSLDSNGEMRSRLSLSEVLDSGDLMKFAVDKTGCQFLEKAVKGSLTSYQKFQLFEQVIGRKDD FLKLSTNIFGNYLVQSVIGISLATNDDGYTKRQEKLKNFISSQMTDMCLDKFACRVIQSSLQNMDLSLAC KLVQALPRDARLIAICVDQNANHVIQKVVAVIPLKNWEFIVDFVATPEHLRQICSDKYGCRVVQTIIEKL TADSMNVDLTSAAQNLRERALQRLMTSVTNRCQELATNEYANYIIQHIVSNDDLAVYRECIIEKCLMRNL LSLSQEKFASHVVEKAFLHAPLELLAEMMDEIFDGYIPHPDTGKDALDIMMFHQFGNYVVQCMLTICCDA VSGRRQTKEGGYDHAISFQDWLKKLHSRVTKERHRLSRFSSGKKMIETLANLRSTHPIYELQ
gi|308387836|pdb|3LTJ|	aRep	5	>gi|308387836|pdb|3LTJ|A Chain A, Structure Of A New Family Of Artificial Alpha Helicoidal Repeat Proteins (Alpha-Rep) Based On Thermostable Heat-Like Repeats MRGSHHHHHHTDPEKVEMYIKNLQDDSYYVRRAAAYALGKIGDERAVEPLIKALKDEDAWVRRAAADALGQIGD ERAVEPLIKALKDEDGWVRQSAAVALGQIGDERAVEPLIKALKDEDWFVRIAAAFALGEIGDERAVEPLIKALKDED GWVRQSAADALGEIGGERVRAAMEKLAETGTGFARKVAVNYLETHKSLIS
gi|109157579|pdb|2FAO|	Pseudomonas Aeruginosa Ligd Polymerase Domain	6	>gi|109157579|pdb|2FAO|A Chain A, Crystal Structure Of Pseudomonas Aeruginosa Ligd Polymerase Domain MGARKASAGASRAATAGVRISHPQRLIDPSIQASKLELAEFHARYADLLLRDLRERPVSLVRGPDGIGGE LFFQKHAARLKIPGIVQLDPALDPGHPPLLQIRSAEALVGAVQMGSIEFHTWNASLANLERPDRFVLDLD PDPALPWKRMLEATQLSLTLLDELGLRAFLKTSGGKGMHLLVPLERRHGWDEVKDFAQAISQHLARLMPE RFSAVSGPRNRVGKIFVDYLRNSRGASTVAAYSVRAREGLPVSVPVFREELDSLQGANQWNLRSLPQRLD ELAGDDPWADYAGTRQRISAAMRRQLGRG
2R9L_A GI: 164519498	Mycobacterium Tuberculosis Ligase D	7	Polymerase Domain From Mycobacterium Tuberculosis Ligase D In Complex With Dna:Accession: 2R9L_A GI: 164519498 >gi|164519498|pdb|2R9L|A Chain A, Polymerase Domain From Mycobacterium Tuberculosis Ligase D In Complex With Dna GSHMGSASEQRVTLTNADKVLYPATGTTKSDIFDYYAGVAEVMLGHIAGRPATRKRWPNGVDQPAFFEKQ LALSAPPWLSRATVAHRSGTTTYPIIDSATGLAWIAQQAALEVHVPQWRFVAEPGSGELNPGPATRLVFD LDPGEGVMMAQLAEVARAVRDLLADIGLVTFPVTSGSKGLHLYTPLDEPVSSRGATVLAKRVAQRLEQAM PALVTSTMTKSLRAGKVFVDWSQNSGSKTTIAPYSLRGRTHPTVAAPRTWAELDDPALRQLSYDEVLTRI ARDGDLLERLDADAPVADRLTRY
1KJZ_A GI:20664108	Large Gamma Subunit Of Initiation Factor Eif2 From Pyrococcus Abyssi	8	Structure Of The Large Gamma Subunit Of Initiation Factor Eif2 From Pyrococcus Abyssi-G235d Mutant:1KJZ_A GI:20664108, >gi|20664108|pdb|1KJZ|A Chain A, Structure Of The Large Gamma Subunit Of Initiation Factor Eif2 From Pyrococcus Abyssi-G235d Mutant GEKRKSRQAEVNIGMVGHVDHGKTTLTKALTGVWTDTHSEELRRGITIKIGFADAEIRRCPNCGRYSTSP VCPYCGHETEFVRRVSFIDAPGHEALMTTMLAGASLMDGAILVIAANEPCPRPQTREHLMALQIIGQKNI IIAQNKIELVDKEKALENYRQIKEFIEGTVAENAPIIPISALHGANIDVLVKAIEDFIPTPKRDPNKPPK MLVLRSFDVNKPGTPPEKLVGGVLDGSIVQGKLKVGDEIEIRPGVPYEEHGRIKYEPITTEIVSLQAGGQ FVEEAYPGGLVGVGTKLDPYLTKGDLMAGNVVGKPGKLPPVWDSLRLEVHLLERVVGTEQELKVEPIKRK EVLLLNVGTARTMGLVTGLGKDEIEVKLQIPVCAEPGDRVAISRQIGSRWRLIGYGIIKE
2D74_A GI:112490420	Translation Initiation Factor Aif2betagamma	9	Crystal Structure Of Translation Initiation Factor Aif2betagamma Heterodimer:2D74_A GI:112490420, >gi|112490420|pdb|2D74|A Chain A, Crystal Structure Of Translation Initiation Factor Aif2betagamma Heterodimer MGEKRKTRQAEVNIGMVGHVDHGKTTLTKALTGVWTDTHSEELRRGITIKIGFADAEIRRCSNCGRYSTS PICPYCGHETEFIRRVSFIDSPGHEALMTTMLAGASLMDGAILVIAANEPCPRPQTREHLMALQIIGQKN IIIAQNKIELVDKEKALENYRQIKEFIKGTVAENAPIIPISALHGANIDVLVKAIEEFIPTPKRDSNKPP

			KMLVLRSFDVNKPGTPPEKLVGGVLDGSIVQGKLKVGDEIEIRPGVPYEEHGRIKYEPITTEIVSLQAGG QFVEEAYPGGLVGIGTKLDPYLTKGDLMAGNVVGKPGKLPPVWTDLRLEVHLLERVVGTEQELNVEPIKR KEVLLLNVGTARTMGLVTALGKDEIELKLQIPVCAEPGERVAISRQIGSRWRLIGYGIIKELEHHHHHH

Table 1: List of enhancer domains for engineered core TALE scaffold.

In another preferred embodiment according to the method of the present invention, the catalytic domain that is capable of processing DNA nearby the single double-stranded DNA target sequence of interest, when fused to said engineered core TALE scaffold according to the method of the present invention, is fused to the N-terminus part of said core TALE scaffold. In another preferred embodiment, said catalytic domain is fused to the C-terminus part of said core TALE scaffold. In another preferred embodiment two catalytic domains are fused to both N-terminus part of said core TALE scaffold and C-terminus part of said core TALE scaffold. In a more preferred embodiment, said catalytic domain has an enzymatic activity selected from the group consisting of nuclease activity, polymerase activity, kinase activity, phosphatase activity, methylase activity, topoisomerase activity, integrase activity, transposase activity or ligase activity. In another preferred embodiment, the catalytic domain fused to the core TALE scaffold of the present invention can be a transcription activator or repressor (i.e. a transcription regulator), or a protein that interacts with or modifies other proteins such as histones. Non-limiting examples of DNA processing activities of said compact TALEN of the present invention include, for example, creating or modifying epigenetic regulatory elements, making site-specific insertions, deletions, or repairs in DNA, controlling gene expression, and modifying chromatin structure.

In another more preferred embodiment, said catalytic domain has an endonuclease activity. In another more preferred embodiment, said catalytic domain has cleavage activity on said double-stranded DNA according to the method of the present invention. In another more preferred embodiment, said catalytic domain has a nickase activity on said double-stranded DNA according to the method of the present invention. In another more preferred embodiment, said catalytic domain is selected from the group consisting of proteins MmeI, Colicin-E7 (CEA7_ECOLX), Colicin-E9, APFL, EndA, Endo I (END1_ECOLI), Human Endo G (NUCG_HUMAN), Bovine Endo G (NUCG_BOVIN), R.HinP1I, I-BasI, I-BmoI, I-HmuI, I-TevI, I-TevII, I-TevIII, I-TwoI, R.MspI, R.MvaI, NucA, NucM, Vvn, Vvn_CLS, Staphylococcal nuclease (NUC_STAAU), Staphylococcal nuclease (NUC_STAHY), Micrococcal nuclease (NUC_SHIFL), Endonuclease yncB, Endodeoxyribonuclease I (ENRN_BPT7), Metnase, Nb.BsrDI, BsrDI A, Nt.BspD6I (R.BspD6I large subunit), ss.BspD6I (R.BspD6I small subunit), R.PleI, MlyI, AlwI, Mva1269I, BsrI, BsmI, Nb.BtsCI, Nt.BtsCI, R1.BtsI, R2.BtsI, BbvCI subunit 1, BbvCI

subunit 2, Bpu10I alpha subunit, Bpu10I beta subunit, BmrI, BfiI, I-CreI, hExoI (EXO1_HUMAN), Yeast ExoI (EXO1_YEAST), E.coli ExoI, Human TREX2, Mouse TREX1, Human TREX1, Bovine TREX1, Rat TREX1, Human DNA2, Yeast DNA2 (DNA2_YEAST) and VP16, as listed in Table 2 (SEQ ID NO: 10 to SEQ ID NO: 66 and SEQ ID NO: 1, 366 & 367), a functional mutant, a variant or a derivative thereof. In another preferred embodiment according to the method of the present invention, said catalytic domain is I-TevI (SEQ ID NO: 20), a functional mutant, a variant or a derivative thereof. In another preferred embodiment, catalytic domain I-TevI (SEQ ID NO: 20), a functional mutant, a variant or a derivative thereof is fused to the N-terminal domain of said core TALE scaffold according to the method of the present invention. In another preferred embodiment, said compact TALEN according to the method of the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group of SEQ ID NO: 420-432.

In another preferred embodiment, said catalytic domain is ColE7 (SEQ ID NO: 11), a functional mutant, a variant or a derivative thereof. In another preferred embodiment, catalytic domain ColE7 (SEQ ID NO: 11), a functional mutant, a variant or a derivative thereof is fused to the N-terminal domain of said core TALE scaffold according to the method of the present invention. In another preferred embodiment, catalytic domain ColE7 (SEQ ID NO: 11), a functional mutant, a variant or a derivative thereof is fused to the C-terminal domain of said core TALE scaffold according to the method of the present invention. In another preferred embodiment, said compact TALEN according to the method of the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group of SEQ ID NO: 435-438.

In another preferred embodiment, said catalytic domain is NucA (SEQ ID NO: 26), a functional mutant, a variant or a derivative thereof. In another preferred embodiment, catalytic domain NucA (SEQ ID NO: 26), a functional mutant, a variant or a derivative thereof is fused to the N-terminal domain of said core TALE scaffold according to the method of the present invention. In another preferred embodiment, catalytic domain NucA (SEQ ID NO: 26), a functional mutant, a variant or a derivative thereof is fused to the C-terminal domain of said core TALE scaffold according to the method of the present invention. In another preferred embodiment, said compact TALEN according to the method of the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group of SEQ ID NO: 433-434.

In another preferred embodiment, said catalytic domain is I-CreI (SEQ ID NO: 1), a functional mutant, a variant or a derivative thereof. In another preferred embodiment, catalytic domain I-CreI (SEQ ID

NO: 1), a functional mutant, a variant or a derivative thereof is fused to the N-terminal domain of said core TALE scaffold according to the method of the present invention. In another preferred embodiment, catalytic domain I-CreI (SEQ ID NO: 1), a functional mutant, a variant or a derivative thereof is fused to the C-terminal domain of said core TALE scaffold according to the method of the present invention. In another preferred embodiment, said compact TALEN according to the method of the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group of SEQ ID NO: 439-441 and SEQ ID NO: 444-446.

In another embodiment, said catalytic domain is a restriction enzyme such as MmeI, R-HinPII, R.MspI, R.MvaI, Nb.BsrDI, BsrDI A, Nt.BspD6I, ss.BspD6I, R.PleI, MlyI and AlwI as non-limiting examples listed in table 2. In another more preferred embodiment, said catalytic domain has an exonuclease activity.

In another more preferred embodiment, any combinations of two catalytic domains selected from the group consisting of proteins MmeI, Colicin-E7 (CEA7_ECOLX), Colicin-E9, APFL, EndA, Endo I (END1_ECOLI), Human Endo G (NUCG_HUMAN), Bovine Endo G (NUCG_BOVIN), R.HinP1I, I-BasI, I-BmoI, I-HmuI, I-TevI, I-TevII, I-TevIII, I-TwoI, R.MspI, R.MvaI, NucA, NucM, Vvn, Vvn_CLS, Staphylococcal nuclease (NUC_STAAU), Staphylococcal nuclease (NUC_STAHY), Micrococcal nuclease (NUC_SHIFL), Endonuclease yncB, Endodeoxyribonuclease I (ENRN_BPT7), Metnase, Nb.BsrDI, BsrDI A, Nt.BspD6I (R.BspD6I large subunit), ss.BspD6I (R.BspD6I small subunit), R.PleI, MlyI, AlwI, Mva1269I, BsrI, BsmI, Nb.BtsCI, Nt.BtsCI, R1.BtsI, R2.BtsI, BbvCI subunit 1, BbvCI subunit 2, Bpu10I alpha subunit, Bpu10I beta subunit, BmrI, BfiI, I-CreI, hExoI (EXO1_HUMAN), Yeast ExoI (EXO1_YEAST), E.coli ExoI, Human TREX2, Mouse TREX1, Human TREX1, Bovine TREX1, Rat TREX1, Human DNA2, Yeast DNA2 (DNA2_YEAST) and VP16, as listed in Table 2 (SEQ ID NO: 10 to SEQ ID NO: 66 and SEQ ID NO: 1, 366 & 367), a functional mutant, a variant or a derivative of these protein domains thereof, can be fused to both N-terminus part and C-terminus part of said core TALE scaffold, respectively. For example, I-HmuI catalytic domain can be fused to the N-terminus part of said core TALE scaffold and ColE7 catalytic domain can be fused to the C-terminus part of said core TALE scaffold. In another example, I-TevI catalytic domain can be fused to the N-terminus part of said core TALE scaffold and ColE7 catalytic domain can be fused to the C-terminus part of said core TALE scaffold. In another embodiment, according to the method of the present invention, said unique compact TALEN monomer comprises a combination of two catalytic domains respectively

fused to the C-terminus part and to the N-terminus part of said core TALE scaffold selected from the group consisting of:

(i) A Nuc A domain (SEQ ID NO: 26) in N-terminus and a Nuc A domain (SEQ ID NO: 26) in C-terminus;

(ii) A ColE7 domain (SEQ ID NO: 11) in N-terminus and a ColE7 domain (SEQ ID NO: 11) in C-terminus;

(iii) A TevI domain (SEQ ID NO: 20) in N-terminus and a ColE7 domain (SEQ ID NO: 11) in C-terminus;

(iv) A TevI domain (SEQ ID NO: 20) in N-terminus and a NucA domain (SEQ ID NO: 26) in C-terminus;

(v) A ColE7 domain (SEQ ID NO: 11) in N-terminus and a NucA domain (SEQ ID NO: 26) in C-terminus;

(vi) A NucA domain (SEQ ID NO: 26) in N-terminus and a ColE7 domain (SEQ ID NO: 11) in C-terminus.

In another preferred embodiment, said compact TALEN according to the method of the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group consisting of SEQ ID NO: 448 and 450.

In another preferred embodiment, said compact TALEN according to the method of the present invention comprises a combination of two catalytic domains respectively fused to the C-terminus part and to the N-terminus part of said core TALE scaffold selected from the group consisting of:

(i) A TevI domain (SEQ ID NO: 20) in N-terminus and a FokI domain (SEQ ID NO: 368) in C-terminus;

(ii) A TevI domain (SEQ ID NO: 20) in N-terminus and a TevI domain (SEQ ID NO: 20) in C-terminus;

(iii) A scTrex2 domain (SEQ ID NO: 451) in N-terminus and a FokI domain (SEQ ID NO: 368) in C-terminus.

In another preferred embodiment, said compact TALEN according to the method of the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group consisting of SEQ ID NO: 447-450 and SEQ ID NO: 452.

In the scope of the present invention, it can be envisioned to insert said catalytic domain between two parts of the engineered core TALE scaffold according to the invention, each part comprising one set of RVDs. In this last case, the number of RVDs for each part of the engineered core TALE scaffold can be the same or not. In other words, it can be envisioned to split said core TALE scaffold of the present invention to insert one catalytic domain between the resulting two parts of said engineered core TALE scaffold. In another preferred embodiment, said compact TALEN according to the method of the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group consisting of SEQ ID NO: 453-455.

GENBANK/SWISS-PROT ID	NAME	SEQ ID NO	FASTA SEQUENCE
ACC85607.1	**MmeI**	10	>gi|186469979|gb|ACC85607.1| MmeI [Methylophilus methylotrophus] MALSWNEIRRKAIEFSKRWEDASDENSQAKPFLIDFFEVFGITNKRVATFEHAVKKFAKAHKEQSRGFVD LFWPGILLIEMKSRGKDLDKAYDQALDYFSGIAERDLPRYVLVCDFQRFRLTDLITKESVEFLLKDLYQN VRSFGFIAGYQTQVIKPQDPINIKAAERMGKLHDTLKLVGYEGHALELYLVRLLFCLFAEDTTIFEKSLF QEYIETKTLEDGSDLAHHINTLFYVLNTPEQKRLKNLDEHLAAFPYINGKLFEEPLPPAQFDKAMREALL DLCSLDWSRISPAIFGSLFQSIMDAKKRRNLGAHYTSEANILKLIKPLFLDELWVEFEKVKNNKNKLLAF HKKLRGLTFFDPACGCGNFLVITYRELRLLEIEVLRGLHRGGQQVLDIEHLIQINVDQFFGIEIEEFPAQ IAQVALWLTDHQMNMKISDEFGNYFARIPLKSTPHILNANALQIDWNDVLEAKKCCFILGNPPFVGKSKQ TPGQKADLLSVFGNLKSASDLDLVAAWYPKAAHYIQTNANIRCAFVSTNSITQGEQVSLLWPLLLSLGIK INFAHRTFSWTNEASGVAAVHCVIIGFGLKDSDEKIIYEYESINGEPLAIKAKNINPYLRDGVDVIACKR QQPISKLPSMRYGNKPTDDGNFLFTDEEKNQFITNEPSSEKYFRRFVGGDEFINNTSRWCLWLDGADISE IRAMPLVLARIKKVQEFRLKSSAKPTRQSASTPMKFFYISQPDTDYLLIPETSSENRQFIPIGFVDRNVI SSNATYHIPSAEPLIFGLLSSTMHNCWMRNVGGRLESRYRYSASLVYNTFPWIQPNEKQSKAIEEAAFAI LKARSNYPNESLAGLYDPKTMPSELLKAHQKLDKAVDSVYGFKGPNTEIARIAFLFETYQKMTSLLPPEK EIKKSKGKN
Q47112.2	**Colicin-E7 (CEA7_ECOLX)**	11	>gi|12644448|sp|Q47112.2|CEA7_ECOLX RecName: Full=Colicin-E7 MSGGDGRGHNSGAHNTGGNINGGPTGLGGNGGASDGSGWSSENNPWGGGSGSGVHWGGGSGHGNGGGNSN SGGGSNSSVAAPMAFGFPALAAPGAGTLGISVSGEALSAAIADIFAALKGPFKFSAWGIALYGILPSEIA KDDPNMMSKIVTSLPAETVTNVQVSTLPLDQATVSVTKRVTDVVKDTRQHIAVVAGVPMSVPVVNAKPTR TPGVFHASFPGVPSLTVSTVKGLPVSTTLPRGITEDKGRTAVPAGFTFGGGSHEAVIRFPKESGQKPVYV SVTDVLTPAQVKQRQDEEKRLQQEWNDAHPVEVAERNYEQARAELNQANKDVARNQERQAKAVQVYNSRK SELDAANKTLADAKAEIKQFERFAREPMAAGHRMWQMAGLKAQRAQTDVNNKKAAFDAAAKEKSDADVAL SSALERRKQKENKEKDAKAKLDKESKRNKPGKATGKGKPVNNKWLNNAGKDLGSPVPDRIANKLRDKEFK SFDDFRKKFWEEVSKDPELSKQFSRNNNDRMKVGKAPKTRTQDVSGKRTSFELHHEKPISQNGGVYDMDN ISVVTPKRHIDIHRGK
CAA38134.1	**EndA**	12	>gi|47374|emb|CAA38134.1| EndA [Streptococcus pneumoniae] MNKKTRQTLIGLLVLLLLSTGSYYIKQMPSAPNSPKTNLSQKKQASEAPSQALAESVLTDAVKSQIKGSL EWNGSGAFIVNGNKTNLDAKVSSKPYADNKTKTVGKETVPTVANALLSKATRQYKNRKETGNGSTSWTPP GWHQVKNLKGSYTHAVDRGHLLGYALIGGLDGFDASTSNPKNIAVQTAWANQAQAEYSTGQNYYESKVRK ALDQNKRVRYRVTLYYASNEDLVPSASQIEAKSSDGELEFNVLVPNVQKGLQLDYRTGEVTVTQ
P25736.1	**Endo I (END1_ECOLI)**	13	>gi|119325|sp|P25736.1|END1_ECOLI RecName: Full=Endonuclease-1; AltName: Full=Endonuclease I; Short=Endo I; Flags: Precursor MYRYLSIAAVVLSAAFSGPALAEGINSFSQAKAAAVKVHADAPGTFYCGCKINWQGKKGVVDLQSCGYQV RKNENRASRVEWEHVVPAWQFGHQRQCWQDGGRKNCAKDPVYRKMESDMHNLQPSVGEVNGDRGNFMYSQ WNGGEGQYGQCAMKVDFKEKAAEPPARARGAIARTYFYMRDQYNLTLSRQQTQLFNAWNKMYPVTDWECE RDERIAKVQGNHNPYVQRACQARKS
Q14249.4	**Human Endo G (NUCG_HUMAN)**	14	>gi|317373579|sp|Q14249.4|NUCG_HUMAN RecName: Full=Endonuclease G, mitochondrial; Short=Endo G; Flags: Precursor MRALRAGLTLASGAGLGAVVEGWRRRREDARAAPGLLGRLPVLPVAAAAELPPVPGGPRGPGELAKYGLP GLAQLKSRESYVLCYDPRTRGALWVVEQLRPERLRGDGDRRECDFREDDSVHAYHRATNADYRGSGFDRG HLAAAANHRWSQKAMDDTFYLSNVAPQVPHLNQNAWNNLEKYSRSLTRSYQNVYVCTGPLFLPRTEADGK SYVKYQVIGKNHVAVPTHFFKVLILEAAGGQIELRTYVMPNAPVDEAIPLERFLVPIESIERASGLLFVP NILARAGSLKAITAGSK
P38447.1	**Bovine Endo G (NUCG_BOVIN)**	15	>gi|585596|sp|P38447.1|NUCG_BOVIN RecName: Full=Endonuclease G, mitochondrial; Short=Endo G; Flags: Precursor MQLLRAGLTLALGAGLGAAAESWWRQRADARATPGLLSRLPVLPVAAAAGLPAVPGAPAGGGPGELAKYG LPGVAQLKSRASYVLCYDPRTRGALWVVEQLRPEGLRGDGNRSSCDFHEDDSVHAYHRATNADYRGSGFD RGHLAAAANHRWSQKAMDDTFYLSNVAPQVPHLNQNAWNNLEKYSRSLTRTYQNVYVCTGPLFLPRTEAD GKSYVKYQVIGKNHVAVPTHFFKVLILEAAGGQIELRSYVMPNAPVDEAIPLEHFLVPIESIERASGLLF VPNILARAGSLKAITAGSK

AAW33811.1	**R.HinP1I**	16	>gi|57116674|gb|AAW33811.1| R.HinP1I restriction endonuclease [Haemophilus influenzae] MNLVELGSKTAKDGFKNEKDIADRFENWKENSEAQDWLVTMGHNLDEIKSVKAVVLSGYKSDINVQVLVF YKDALDIHNIQVKLVSNKRGFNQIDKHWLAHYQEMWKFDDNLLRILRHFTGELPPYHSNTKDKRRMFMTE FSQEEQNIVLNWLEKNRVLVLTDILRGRGDFAAEWVLVAQKVSNNARWILRNINEVLQHYGSGDISLSPR GSINFGRVTIQRKGGDNGRETANMLQFKIDPTELFDI
AAO93095.1	**I-BasI**	17	>gi|29838473|gb|AAO93095.1| I-BasI [Bacillus phage Bastille] MFQEEWKDVTGFEDYYEVSNKGRVASKRTGVIMAQYKINSGYLCIKFTVNKKRTSHLVHRLVAREFCEGY SPELDVNHKDTDRMNNNYDNLEWLTRADNLKDVRERGKLNTHTAREALAKVSKKAVDVYTKDGSEYIATY PSATEAAEALGVQGAKISTVCHGKRQHTGGYHFKFNSSVDPNRSVSKK
AAK09365.1	**I-BmoI**	18	>gi|12958590|gb|AAK09365.1|AF321518_2 intron encoded I-BmoI [Bacillus mojavensis] MKSGVYKITNKNTGKFYIGSSEDCESRLKVHFRNLKNNRHINRYLNNSFNKHGEQVFIGEVIHILPIEEA IAKEQWYIDNFYEEMYNISKSAYHGGDLTSYHPDKRNIILKRADSLKKVYLKMTSEEKAKRWQCVQGENN PMFGRKHTETTKLKISNHNKLYYSTHKNPFKGKKHSEESKTKLSEYASQRVGEKNPFYGKTHSDEFKTYM SKKFKGRKPKNSRPVIIDGTEYESATEASRQLNVVPATILHRIKSKNEKYSGYFYK
P34081.1	**I-HmuI**	19	>gi|465641|sp|P34081.1|HMUI_BPSP1 RecName: Full=DNA endonuclease I-HmuI; AltName: Full=HNH homing endonuclease I-HmuI MEWKDIKGYEGHYQVSNTGEVYSIKSGKTLKHQIPKDGYHRIGLFKGGKGKTFQVHRLVAIHFCEGYEEG LVVDHKDGNKDNNLSTNLRWVTQKINVENQMSRGTLNVSKAQQIAKIKNQKPIIVISPDGIEKEYPSTKC ACEELGLTRGKVTDVLKGHRIHHKGYTFRYKLNG
P13299.2	**I-TevI**	20	>gi|6094464|sp|P13299.2|TEV1_BPT4 RecName: Full=Intron-associated endonuclease 1; AltName: Full=I-TevI; AltName: Full=IRF protein MKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDL IIERENFWIKELNSKINGYNIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKN GRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSFFNHKHSDITKSKISEKMKGKKPSNIKKISCDGV IFDCAADAARHFKISSGLVTYRVKSDKWNWFYINA
P07072.2	**I-TevII**	21	>gi|20141823|sp|P07072.2|TEV2_BPT4 RecName: Full=Intron-associated endonuclease 2; AltName: Full=I-TevII MKWKLRKSLKIANSVAFTYMVRFPDKSFYIGFKKFKTIYGKDTNWKEYNSSSKLVKEKLKDYKAKWIILQ VFDSYESALKHEEMLIRKYFNNEFILNKSIGGYKFNKYPDSEEHKQKLSNAHKGKILSLKHKDKIREKLI EHYKNNSRSEAHVKNNIGSRTAKKTVSIALKSGNKFRSFKSAAKFLKCSEEQVSNHPNVIDIKITIHPVP EYVKINDNIYKSFVDAAKDLKLHPSRIKDLCLDDNYPNYIVSYKRVEK
Q38419.1	**I-TevIII**	22	>gi|11387192|sp|Q38419.1|TEV3_BPR03 RecName: Full=Intron-associated endonuclease 3; AltName: Full=I-TevIII MNYRKIWIDANGPIPKDSDGRTDEIHHKDGNRENNDLDNLMCLSIQEHYDIHLAQKDYQACHAIKLRMKY SPEEISELASKAAKSREIQIFNIPEVRAKNIASIKSKIENGTFHLLDGEIQRKSNLNRVALGIHNFQQAE HIAKVKERNIAAIKEGTHVFCGGKMQSETQSKRVNDGSHHFLSEDHKKRTSAKTLEMVKNGTHPAQKEIT CDFCGHIGKGPGFYLKHNDRCKLNPNRIQLNCPYCDKKDLSPSTYKRWHGDNCKARFND
AAM00817.1	**I-TwoI**	23	>gi|19881200|gb|AAM00817.1|AF485080_2 HNH endonuclease I-TwoI [Staphylococcus phage Twort] MEELWKEIPGFNSYMISNKGQVYSRKRNKILALRTDKNGYKRISIFNNEGKRILLGVHKLVLLGFKGINT EKPIPHHKNNIKDDNRLENLEWVTVSENTKHAYDIGALKSPRRVTCTLYYKGEPLSCYDSLFDLAKALKV SRSVIESPRNGLVLSTFEVKREPTIQGLPLNKEIFEHSLIKGLGNPPLKVYNEDETYYFLTLMDISKYFN ESYSKVQRGYYKGKWKSYIIEHIDFYEYYKQTH
P11405.1	**R.MspI**	24	>gi|135239|sp|P11405.1|T2M1_MORSP RecName: Full=Type-2 restriction enzyme MspI; Short=R.MspI; AltName: Full=Endonuclease MspI; AltName: Full=Type II restriction enzyme MspI MRTELLSKLYDDFGIDQLPHTQHGVTSDRLGKLYEKYILDIFKDIESLKKYNTNAFPQEKDISSKLLKAL NLDLDNIIDVSSSDTDLGRTIAGGSPKTDATIRFTFHNQSSRLVPLNIKHSSKKKVSIAEYDVETICTGV GISDGELKELIRKHQNDQSAKLFTPVQKQRLTELLEPYRERFIRWCVTLRAEKSEGNILHPDLLIRFQVI DREYVDVTIKNIDDYVSDRIAEGSKARKPGFGTGLNWTYASGSKAKKMQFKG
R.MvaI	**R.MvaI**	25	>gi|119392963|gb|AAM03024.2|AF472612_1 R.MvaI [Kocuria varians] MSEYLNLLKEAIQNVVDGGWHETKRKGNTGIGKTFEDLLEKEEDNLDAPDFHDIEIKTHETAAKSLLTLF TKSPTNPRGANTMLRNRYGKKDEYGNNILHQTVSGNRKTNSNSYNYDFKIDIDWESQVVRLEVFDKQDIM IDNSVYWSFDSLQNQLDKKLKYIAVISAESKIENEKKYYKYNSANLFTDLTVQSLCRGIENGDIKVDIRI GAYHSGKKKGKTHDHGTAFRINMEKLLEYGEVKVIV
CAA45962.1	**NucA**	26	>gi|39041|emb|CAA45962.1| NucA [Nostoc sp. PCC 7120] MGICGKLGVAALVALIVGCSPVQSQVPPLTELSPSISVHLLLGNPSGATPTKLTPDNYLMVKNQYALSYN NSKGTANWVAWQLNSSWLGNAERQDNFRPDKTLPAGWVRVTPSMYSGSGYDRGHIAPSADRTKTTEDNAA TFLMTNMMPQTPDNNRNTWGNLEDYCRELVSQGKELYIVAGPNGSLGKPLKGKVTVPKSTWKIVVVLDSP GSGLEGITANTRVIAVNIPNDPELNNDWRAYKVSVDELESLTGYDFLSNVSPNIQTSIESKVDN
P37994.2	**NucM**	27	>gi|313104150|sp|P37994.2|NUCM_DICD3 RecName: Full=Nuclease nucM; Flags: Precursor MLRNLVIFAVLGAGLTTLAAAGQDINNFTQAKAAAAKIHQDAPGTFYCGCKINWQGKKGTPDLASCGYQV RKDANRASRIEWEHVVPAWQFGHQRQCWQDGGRKNCTKDDVYRQIETDLHNLQPAIGEVNGDRGNFMYSQ WNGGERQYGQCEMKIDFKSQLAEPPERARGAIARTYFYMRDRYNLNLSRQQTQLFDAWNKQYPATTWECT REKRIAAVQGNHNPYVQQACQP
AAF19759.1	**Vvn**	28	>gi|6635279|gb|AAF19759.1|AF063303_1 nuclease precursor Vvn [Vibrio vulnificus] MKRLFIFIASFTAFAIQAAPPSSFSAAKQQAVKIYQDHPISFYCGCDIEWQGKKGIPNLETCGYQVRKQQ TRASRIEWEHVVPAWQFGHHRQCWQKGGRKNCSKNDQQFRLMEADLHNLTPAIGEVNGDRSNFNFSQWNG VDGVSYGRCEMQVNFKQRKVMPQTELRGSIARTYLYMSQEYGFQLSKQQQQLMQAWNKSYPVDEWECTRD

			DRIAKIQGNHNPFVQQSCQTQ
AAF19759.1 (reference)	**Vvn_CLS**	29	>Vvn_CLS (variant of AAF19759.1) MASGAPPSSFSAAKQQAVKIYQDHPISFYCGCDIEWQGKKGIPNLETCGYQVRKQQTRASRIEWEHVVPA WQFGHHRQCWQKGGRKNCSKNDQQFRLMEADLHNLTPAIGEVNGDRSNFNFSQWNGVDGVSYGRCEMQVN FKQRKVMPPDRARGSIARTYLYMSQEYGFQLSKQQQQLMQAWNKSYPVDEWECTRDDRIAKIQGNHNPFV QQSCQTQGSSAD
P00644.1	**Staphylococcal nuclease (NUC_STAAU)**	30	>gi|128852|sp|P00644.1|NUC_STAAU RecName: Full=Thermonuclease; Short=TNase; AltName: Full=Micrococcal nuclease; AltName: Full=Staphylococcal nuclease; Contains: RecName: Full=Nuclease B; Contains: RecName: Full=Nuclease A; Flags: Precursor MLVMTEYLLSAGICMAIVSILLIGMAISNVSKGQYAKRFFFFATSCLVLTLVVVSSLSSSANASQTDNGV NRSGSEDPTVYSATSTKKLHKEPATLIKAIDGDTVKLMYKGQPMTFRLLLVDTPETKHPKKGVEKYGPEA SAFTKKMVENAKKIEVEFDKGQRTDKYGRGLAYIYADGKMVNEALVRQGLAKVAYVYKPNNTHEQHLRKS EAQAKKEKLNIWSEDNADSGQ
P43270.1	**Staphylococcal nuclease (NUC_STAHY)**	31	>gi|1171859|sp|P43270.1|NUC_STAHY RecName: Full=Thermonuclease; Short=TNase; AltName: Full=Micrococcal nuclease; AltName: Full=Staphylococcal nuclease; Flags: Precursor MKKITTGLIIVVAAIIVLSIQFMTESGPFKSAGLSNANEQTYKVIRVIDGDTIIVDKDGKQQNLRMIGVD TPETVKPNTPVQPYGKEASDFTKRHLTNQKVRLEYDKQEKDRYGRTLAYVWLGKEMFNEKLAKEGLARAK FYRPNYKYQERIEQAQKQAQKLKKNIWSN
P29769.1	**Micrococcal nuclease (NUC_SHIFL)**	32	>gi|266681|sp|P29769.1|NUC_SHIFL RecName: Full=Micrococcal nuclease; Flags: Precursor MKSALAALRAVAAAVVLIVSVPAWADFRGEVVRILDGDTIDVLVNRQTIRVRLADIDAPESGQAFGSRAR QRLADLTFRQEVQVTEKEVDRYGRTLGVVYAPLQYPGGQTQLTNINAIMVQEGMAWAYRYYGKPTDAQMY EYEKEARRQRLGLWSDPNAQEPWKWRRASKNATN
P94492.1	**Endonuclease yncB**	33	>gi|81345826|sp||YNCB_BACSU RecName: Full=Endonuclease yncB; Flags: Precursor MKKILISMIAIVLSITLAACGSNHAAKNHSDSNGTEQVSQDTHSNEYNQTEQKAGTPHSKNQKKLVNVTL DRAIDGDTIKVIYNGKKDTVRYLLVDTPETKKPNSCVQPYGEDASKRNKELVNSGKLQLEFDKGDRRDKY GRLLAYVYVDGKSVQETLLKEGLARVAYVYEPNTKYIDQFRLDEQEAKSDKLSIWSKSGYVTNRGFNGCV K
P00641.1	**Endodeoxyribonuclease I (ENRN_BPT7)**	34	>gi|119370|sp|P00641.1|ENRN_BPT7 RecName: Full=Endodeoxyribonuclease 1; AltName: Full=Endodeoxyribonuclease I; Short=Endonuclease MAGYGAKGIRKVGAFRSGLEDKVSKQLESKGIKFEYEEWKVPYVIPASNHTYTPDFLLPNGIFVETKGLW ESDDRKKHLLIREQHPELDIRIVFSSSRTKLYKGSPTSYGEFCEKHGIKFADKLIPAEWIKEPKKEVPFD RLKRKGGKK
Q53H47.1	**Metnase**	35	>gi|74740552|sp|Q53H47.1|SETMR_HUMAN RecName: Full=Histone-lysine N-methyltransferase SETMAR; AltName: Full=SET domain and mariner transposase fusion gene-containing protein; Short=HsMar1; Short=Metnase; Includes: RecName: Full=Histone-lysine N-methyltransferase; Includes: RecName: Full=Mariner transposase Hsmar1 MAEFKEKPEAPTEQLDVACGQENLPVGAWPPGAAPAPFQYTPDHVVGPGADIDPTQITFPGCICVKTPCL PGTCSCLRHGENYDDNSCLRDIGSGGKYAEPVFECNVLCRCSDHCRNRVVQKGLQFHFQVFKTHKKGWGL RTLEFIPKGRFVCEYAGEVLGFSEVQRRIHLQTKSDSNYIIAIREHVYNGQVMETFVDPTYIGNIGRFLN HSCEPNLLMIPVRIDSMVPKLALFAAKDIVPEEELSYDYSGRYLNLTVSEDKERLDHGKLRKPCYCGAKS CTAFLPFDSSLYCPVEKSNISCGNEKEPSMCGSAPSVFPSCKRLTLETMKMMLDKKQIRAIFLFEFKMGR KAAETTRNINNAFGPGTANERTVQWWFKKFCKGDESLEDEERSGRPSEVDNDQLRAIIEADPLTTTREVA EELNVNHSTVVRHLKQIGKVKKLDKWVPHELTENQKNRRFEVSSSLILRNHNEPFLDRIVTCDEKWILYD NRRRSAQWLDQEEAPKHFPKPILHPKKVMVTIWWSAAGLIHYSFLNPGETITSEKYAQEIDEMNQKLQRL QLALVNRKGPILLHDNARPHVAQPTLQKLNELGYEVLPHPPYSPDLLPTNYHVFKHLNNFLQGKRFHNQQ DAENAFQEFVESQSTDFYATGINQLISRWQKCVDCNGSYFD
ABD15132.1	**Nb.BsrDI**	36	>gi|86757493|gb|ABD15132.1| Nb.BsrDI [Geobacillus stearothermophilus] MTEYDLHLYADSFHEGHWCCENLAKIAQSDGGKHQIDYLQGFIPRHSLIFSDLIINITVFGSYKSWKHLP KQIKDLLFWGKPDFIAYDPKNDKILFAVEETGAVPTGNQALQRCERIYGSARKQIPFWYLLSEFGQHKDG GTRRDSIWPTIMGLKLTQLVKTPSIILHYSDINNPEDYNSGNGLKFLFKSLLQIIINYCTLKNPLKGMLE LLSIQYENMLEFIKSQWKEQIDFLPGEEILNTKTKELARMYASLAIGQTVKIPEELFNWPRTDKVNFKSP QGLIKYDELCYQLEKAVGSKKAYCLSNNAGAKPQKLESLKEWINSQKKLFDKAPKLTPPAEFNMKLDAFP VTSNNNYYVTTSKNILYLFDYWKDLRIAIETAFPRLKGKLPTDIDEKPALIYICNSVKPGRLFGDPFTGQ LSAFSTIFGKKNIDMPRIVVAYYPHQIYSQALPKNNKSNKGITLKKELTDFLIFHGGVVVKLNEGKAY
ABD15133.1	**BsrDI A**	37	>gi|86757494|gb|ABD15133.1| BsrDI A [Geobacillus stearothermophilus] MTDYRYSFELSEEIARWAPEIKTKNTDWFVAFSNPTAGPWKRVMAIDKASNREGEVHRFGREDERPDIIL VNDNISLILILEAKEKLNQLISKSQVDKSVDVFLTLSSILKEKSDNNYWGDRTKYINVLGILWGSEQETS QKDIDNAFRVYRDSLVKNLKEINPTPTNICTDILVGVESIKNKKEEISIKIHVSNIYAEIYPKFTGKHLL EKLAVLN
ABN42182.1	**Nt.BspD6I (R.BspD6I large subunit)**	38	>gi|125396996|gb|ABN42182.1| heterodimeric restriction endonuclease R.BspD6I large subunit [Bacillus sp. D6] MAKKVNWYVSCSPRSPEKIQPELKVLANFEGSYWKGVKGYKAQEAFAKELAALPQFLGTTYKKEAAFSTR DRVAPMKTYGFVFVDEEGYLRITEAGKMLANNRRPKDVFLKQLVKWQYPSFQHKGKEYPEEEWSINPLVF VLSLLKKVGGLSKLDIAMFCLTATNNNQVDEIAEEIMQFRNEREKIKGQNKKLEFTENYFFKRFEKIYGN VGKIREGKSDSSHKSKIETKMRNARDVADATTRYFRYTGLFVARGNQLVLNPEKSDLIDEIISSSKVVKN YTRVEEFHEYYGNPSLPQFSFETKEQLLDLAHRIRDENTRLAEQLVEHFPNVKVEIQVLEDIYNSLNKKV DVETLKDVIYHAKELQLELKKKKLQADFNDPRQLEEVIDLLEVYHEKKNVIEEKIKARFIANKNTVFEWL TWNGFIILGNALEYKNNFVIDEELQPVTHAAGNQPDMEIIYEDFIVLGEVTTSKGATQFKMESEPVTRHY LNKKKELEKQGVEKELYCLFIAPEINKNTFEEFMKYNIVQNTRIIPLSLKQFNMLLMVQKKLIEKGRRLS SYDIKNLMVSLYRTTIECERKYTQIKAGLEETLNNWVVDKEVRF

ABN42183.1	**ss.BspD6I (R.BspD6I small subunit)**	39	>gi|125396997|gb|ABN42183.1| heterodimeric restriction endonuclease R.BspD6I small subunit [Bacillus sp. D6] MQDILDFYEEVEKTINPPNYFEWNTYRVFKKLGSYKNLVPNFKLDDSGHPIGNAIPGVEDILVEYEHFSI LIECSLTIGEKQLDYEGDSVVRHLQEYKKKGIEAYTLFLGKSIDLSFARHIGFNKESEPVIPLTVDQFKK LVTQLKGDGEHFNPNKLKEILIKLLRSDLGYDQAEEWLTFIEYNLK
AAK27215.1	**R.PleI**	40	>gi|13448813|gb|AAK27215.1|AF355461_2 restriction endonuclease R.PleI [Paucimonas lemoignei] MAKPIDSKVLFITTSPRTPEKMVPEIELLDKNFNGDVWNKDTQTAFMKILKEESFFDGEGKNDPAFSARD RINRAPKSLGFVILTPKLSLTDAGVELIKAKRKDDIFLRQMLKFQLPSPYHKLSDKAALFYVKPYLEIFR LVRHFGSLTFDELMIFGLQIIDFRIFNQIVDKIEDFRVGKIENKGRYKTYKKERFEEELGKIYKDELFGL TEASAKTLITKKGNNMRDYADACVRYLRATGMVNVSYQGKSLSIVQEKKEEVDFFLKNTEREPCFINDEA SYVSYLGNPNYPKLFVDDVDRIKKKLRFDFKKTNKVNALTLPELKEELENEILSRKENILKSQISDIKNF KLYEDIQEVFEKIENDRTLSDAPLMLEWNTWRAMTMLDGGEIKANLKFDDFGSPMSTAIGNMPDIVCEYD DFQLSVEVTMASGQKQYEMEGEPVSRHLGKLKKSSEKPVYCLFIAPKINPSSVAHFFMSHKVDIEYYGGK SLIIPLELSVFRKMIEDTFKASYIPKSDNVHKLFKNFASIADEAGNEKVWYEGVKRTAMNWLSLS
AAK39546.1	**MlyI**	41	>gi|13786046|gb|AAK39546.1|AF355462_2 MlyIR [Micrococcus lylae] MASLSKTKHLFGFTSPRTIEKIIPELDILSQQFSGKVWGENQINFFDAIFNSDFYEGTTYPQDPALAARD RITRAPKALGFIQLKPVIQLTKAGNQLVNQKRLPELFTKQLLKFQLPSPYHTQSPTVNFNVRPYLELLRL INELGSISKTEIALFFLQLVNYNKFDEIKNKILKFRETRKNNRSVSWKTYVSQEFEKQISIIFADEVTAK NFRTRESSDESFKKFVKTKEGNMKDYADAFFRYIRGTQLVTIDKNLHLKISSLKQDSVDFLLKNTDRNAL NLSLMEYENYLFDPDQLIVLEDNSGLINSKIKQLDDSINVESLKIDDAKDLLNDLEIQRKAKTIEDTVNH LKLRSDIEDILDVFAKIKKRDVPDVPLFLEWNIWRAFAALNHTQAIEGNFIVDLDGMPLNTAPGKKPDIE INYGSFSCIVEVTMSSGETQFNMEGSSVPRHYGDLVRKVDHDAYCIFIAPKVAPGTKAHFFNLNRLSTKH YGGKTKIIPMSLDDFICFLQVGITHNFQDINKLKNWLDNLINFNLESEDEEIWFEEIISKISTWAI
YP_004134094.1	**AlwI**	42	>gi|319768594|ref|YP_004134094.1| restriction endonuclease, type II, AlwI [Geobacillus sp. Y412MC52] MNKKNTRKVWFITRPERDPRFHQEALLALQKATDDFRLKWAGNREVHKRYEEELANMGIKRNNVSHDGSG GRTWMAMLKTFSYCYVDDDGYIRLTKVGEKLIQGEKVYENTRKQVLTLQYPNAYFLEPGFRPKFDEGFRI RPVLFLIKLANDERLDFYVTKEEITYFAMTAQKDSQLDEIVHKILAFRKAGPREREEMKQDIAAKFDHRE RSDKGARDFYEAHSDVAHTFMLISDYTGLVEYIRGKALKGDSSKINEIKQEIAEIEKRYPFNTRYMISLE RMAENSGLDVDSYKASRYGNIKPAANSSKLRAKAERILAQFPSIESMSKEEIAGALQKYLSPRDIEKVIH EIVENKDDFEGINSDFVETYLNEKDNLAFEDKTGQIFSALGFDVAMRPKAKNGERTEIEIIARYGGSKFG IIDAKNYAGKFPLSSSLVSHMASEYIPNYTGYEGKELTFFGYVTANDFSGERNLEKISDKAKRITGNPIS GFLVTARTLLGFLDYCIENDVPLEDRAELFVKAVKNKGYKSLEALLRELKETI
AAY97906.1	**Mva1269I**	43	>gi|68480350|gb|AAY97906.1| Mva1269I restriction endonuclease [Kocuria varians] MYLNTAVFNIYGDNIVECSRAFHYILEGFKLANISITQEYDLQNITTPKFCIYTDKFRYIFIFIPGTSAS RWNKDIYKELVLNNGGPLKEGADAIITRIFSEDSELVLASMEFSAALPAGNNTWQRSGRAYSLTAANIPY FYIVQLGGKEIKKGKDGKSDKFATRLPNPALSLSFTLNTIKKPAPSLIVYDQAPEADSAISDLYSNCYGI DDFSLYLFKLITEENNLHELKNIYNKNVEFLQLRSVDEKGKNFSGKDYKYIFEHKDPYKGLTEVVKERKI PWKKKTATKTFENFPLRNQAPIFRLIDFLSTKSYGIVSKDSLPLTFIPSEHRVEVANYICNQLYIDKVSD EFVKWIYKKEDLAICIINGFKPGGDDSRPDRGLPPFTKMLTNLDILTLMFGPAPPTQWDYLDSDPEKLNK TNGLWQSIFAFSDAILVDSSTRDNNKFVYNAYLKEHWVVQREKKESNTPISYFPKSVGEHDVDTSLHILF TYIGKHFESACNPPGGDWSGVSLLKNNIEYRWTSMYRVSQDGTKRPDHIYQLVYNSTDTLLLIESKGIKN DLLKSKEANVGIGMINYLKNLMARDYTAVKKDGEWKNIHGQMTLDKFLTFSAVAYLFTTDFDNEYTSAAE LLVHSNTQLAFALEIKEKNSVMHIFTANTVAYNFAEYLLETMRNSHLPLKIYKPI
ADR72996.1	**BsrI**	44	>gi|313667100|gb|ADR72996.1| BsrI [Geobacillus stearothermophilus] MRNIRIYSEVKEQGIFFKEVIQSVLEKANVEVVLVNSAMLDYSDVSVISLIRNQKKFDLLVSEVRDKREI PIVMVEFSTAVTTDDHELQRADAMFWAYKYKIPYLKISPMEKKSQTADDKFGGGRLLSVNDQIIHMYRTD GVMYHIEWESMDNSAYVKNAELYPSCPDCAPELASLFRCLLETIEKCENIEDYYRILLDKLGKQKVAVKW GNFREEKTLEQWKHEKFDLLERFSKSSSRMEYDKDKKELKIKVNRYGHAMDPERGILAFWKLVLGDEWKI VAEFQLQRKTLKGRQSYQSLFDEVSQEEKLMNIASEIIKNGNVISPDKAIEIHKLATSSTMISTIDLGTP ERKYITDDSLKGYLQHGLITNIYKNLLYYVDEIRFTDLQRKTIASLTWNKEIVNDYYKSLMDQLLDKNLR VLPLTSIKNISEDLITWSSKEILINLGYKILAASYPEAQGDRCILVGPTGKKTERKFIDLIAISPKSKGV ILLECKDKLSKSKDDCEKMNDLLNHNYDKVTKLINVLNINNYNYNNIIYTGVAGLIGRKNVDNLPVDFVI KFKYDAKNLKLNWEINSDILGKHSGSFSMEDVAVVRKRS
AAL86024.1	**BsmI**	45	>gi|19347662|gb|AAL86024.1| BsmI [Geobacillus stearothermophilus] MNVFRIHGDNIIECERVIDLILSKINPQKVKRGFISLSCPFIEIIFKEGHDYFHWRFDMFPGFNKNTNDR WNSNILDLLSQKGSFLYETPDVIITSLNNGKEEILMAIEFCSALQAGNQAWQRSGRAYSVGRTGYPYIYI VDFVKYELNNSDRSRKNLRFPNPAIPYSYISHSKNTGNFIVQAYFRGEEYQPKYDKKLKFFDETIFAEDD IADYIIAKLQHRDTSNIEQLLINKNLKMVEFLSKNTKNDNNFTYSEWESIYNGTYRITNLPSLGRFKFRK KIAEKSLSGKVKEFNNIVQRYSVGLASSDLPFGVIRKESRNDFINDVCKLYNINDMKIIKELKEDADLIV CMLKGFKPRGDDNRPDRGALPLVAMLAGENAQIFTFIYGPLIKGAINLIDQDINKLAKRNGLWKSFVSLS DFIVLDCPIIGESYNEFRLIINKNNKESILRKTSKQQNILVDPTPNHYQENDVDTVIYSIFKYIVPNCFS GMCNPPGGDWSGLSIIRNGHEFRWLSLPRVSENGKRPDHVIQILDLFEKPLLLSIESKEKPNDLEPKIGV QLIKYIEYLFDFTPSVQRKIAGGNWEFGNKSLVPNDFILLSAGAFIDYDNLTENDYEKIFEVTGCDLLIA IKNQNNPQKWVIKFKPKNTIAEKLVNYIKLNFKSNIFDTGFFHIEG
ADI24225.1	**Nb.BtsCI**	46	>gi|297185870|gb|ADI24225.1| BtsCI bottom-strand nicking enzyme variant [synthetic construct] MKRILYLLTEERPKINIIHQIINLEYKATLHFGAKIVPVMNEENKFTFIYHVKGIEVEGFDAVLIKIVSG HSSFVDYLVFDSNDLKPEKNTITLFDLDQYELDLSYYFGKGWIVRIPSPSDLPKYVVEETKTDDHESRNT NAYQRSSKFVFCELYYGKEVKKYMLYDISDGRTLSGTDTHNFGMRMLVTNNVNLVGVPNMYLPFTDIKEF INEKNRIADNGPSHNVPIRLKLDKEKNVIYISAKLDKGNGKNKNKISNDPNIGAVAIISATLRNLNWKGD IEIINHNLLPSSISSRSNGNKLLYIMKKLGVRFNNINVNWNNIKNNINYFFYNITSEKIVSIYYHLYVED KLSNARVIFDNHAGCGKSYFRTLNNKIIPVGKEIPLPALVIFDSDQNIVKVIAAAKAENVYNGVEQLSTF DKFIESYINKYYPGAAVECSVITWGKSSNPYVSFYLDKDGSAVFL
ADI24224.1	**Nt.BtsCI**	47	>gi|297185868|gb|ADI24224.1| BtsCI top-strand nicking enzyme variant

			[synthetic construct] MKRILYLLTEERPKINIIHQIINLEYKATLHFGAKIVPVMNEENKFTFIYHVKGIEVEGFDAVLIKIVSG HSSFVDYLVFDSNDLKPEKNTITLFDLDQYELDLSYYFGKGWIVRIPSPSDLPKYVVFETKTDDHESRNT NAYQRSSKFVFCELYYGKEVKKYMLYDISDGRTLSGTDTHNFGMRMLVTNNVNLVGVPNMYLPFTDIKEF INEKNRIADNGPSHNVPIRLKLDKEKNVIYISAKLDKGNGKNKNKISNDPNIGAVAIISATLRNLNWKGD IEIINHNLLPSSISSRSNGNKLLYIMKKLGVRFNNINVNWNNIKNNINYFFYNITSEKIVSIYYHLYVED KLSNARVIFDNHAGCGKSYFRTLNNKIIPVGKEIPLPDLVIFDSDQNIVKVIEAEKAENVYNGVEQLSTF DKFIESYINKYYPGAAVECSVITWGKSSNPYVSFYLDKDGSAVFL >gi|85720924|gb|ABC75874.1| R1.BtsI [Geobacillus thermoglucosidasius] MKITEGIVHVAMRHFLKSNGWKLIAGQYPGGSDDELTALNIVDPVVARDNSPDPRRHSLGKIVPDLIAYK NDDLLVIEAKPKYSQDDRDKLLYLLSERKHDFYAALEKFATERNHPELLPVSKLNIIPGLAFSASENKFK KDPGFVYIRVSGIFEAFMEGYDWG
ABC75874.1	**R1.BtsI**	48	>gi|85720924|gb|ABC75874.1| R1.BtsI [Geobacillus thermoglucosidasius] MKITEGIVHVAMRHFLKSNGWKLIAGQYPGGSDDELTALNIVDPVVARDNSPDPRRHSLGKIVPDLIAYK NDDLLVIEAKPKYSQDDRDKLLYLLSERKHDFYAALEKFATERNHPELLPVSKLNIIPGLAFSASENKFK KDPGFVYIRVSGIFEAFMEGYDWG
ABC75876.1	**R2.BtsI**	49	>gi|85720926|gb|ABC75876.1| R2.BtsI [Geobacillus thermoglucosidasius] MQIEQLMKSLTIYFDDIQEGLWFKNLHPLLESASLEAITGSLKRNPNLADVLKYDRPDIILTLNQTPILV IERTIEVPSGHNVGQRYGRLAAASEAGVPLVYFGPYAARKHGGATEGPRYMNLRLFYALDVMQKVNGSAI TTINWPVDQNFEILQDPSKDKRMKEYLEMFFDNLLKYGIAGINLAIRNSSFQAEQLAEREKFVETMITNP EQYDVPPDSVQILNAERFFNELGISENKRIICDEVVLYQVGMTYVRSDPYTGMALLYKYLYILGSERNRC LILKFPNITTDMWKKVAFGSRERKDVRIYRSVSDGILFADGYLSKEEL
AAX14652.1	**BbvCI subunit 1**	50	>gi|60202520|gb|AAX14652.1| BbvCI endonuclease subunit 1 [Brevibacillus brevis] MINEDFFIYEQLSHKKNLEQKGKNAFDEETEELVRQAKSGYHAFIEGINYDEVTKLDLNSSVAALEDYIS IAKEIEKKHKMFNWRSDYAGSIIPEFLYRIVHVATVKAGLKPIFSTRNTIIEISGAAHREGLQIRRKNED FALGFHEVDVKIASESHRVISLAVACEVKTNIDKNKLNGLDFSAERMKRTYPGSAYFLITETLDFSPDEN HSSGLIDEIYVLRKQVRTKNRVQKAPLCPSVFAELLEDILEISYRASNVKGHVYDRLEGGKLIRV
AAX14653.1	**BbvCI subunit 2**	51	>gi|60202521|gb|AAX14653.1| BbvCI endonuclease subunit 2 [Brevibacillus brevis] MFNQFNPLVYTHGGKLERKSKKDKTASKVFEEFGVMEAYNCWKEASLCIQQRDKDSVLKLVAALNTYKDA VEPIFDSRLNSAQEVLQPSILEEFFEYLFSRIDSIVGVNIPIRHPAKGYLSLSFNPHNIETLIQSPEYTV RAKDHDFIIGGSAKLTIQGHGGEGETTNIVVPAVAIECKRYLERNMLDECAGTAERLKRATPYCLYFVVA EYLKLDDGAPELTEIDEIYILRHQRNSERNKPGFKPNPIDGELIWDLYQEVMNHLGKIWWDPNSALQRGK VFNRP
CAA74998.1	**Bpu10I alpha subunit**	52	>gi|2894388|emb|CAA74998.1| Bpu10I restriction endonuclease alpha subunit [Bacillus pumilus] MGVEQEWIKNITDMYQSPELIPSHASNLLHQLKREKRNEKLKKALEIITPNYISYISILLNNHNMTRKEI VILVDALNEYMNTLRHPSVKSVFSHQADFYSSVLPEFFNLLFRNLIKGLNEKIKVNSQKDIIIDCIFDPY NEGRVVFKKKRVDVAIILKNKFVFNNVEISDFAIPLVAIEIKTNLDKNMLSGIEQSVDSLKETFPLCLYY CITELADFAIEKQNYASTHIDEVFILRKQKRGPVRRGTPLEVVHADLILEVVEQVGEHLSKFKDPIKTLK ARMTEGYLIKGKGK
CAA74999.1	**Bpu10I beta subunit**	53	>gi|2894389|emb|CAA74999.1| Bpu10I restriction endonuclease beta subunit [Bacillus pumilus] MTQIDLSNTKHGSILFEKQKNVKEKYLQQAYKHYLYFRRSIDGLEITNDEAIFKLTQAANNYRDNVLYLF ESRPNSGQEAFRYTILEEFFYHLFKDLVKKKFNQEPSSIVMGKANSYVSLSFSPESFLGLYENPIPYIHT KDQDFVLGCAVDLKISPKNELNKENETEIVVPVIAIECKTYIERNMLDSCAATASRLKAAMPYCLYIVAS EYMKMDQAYPELTDIDEVFILCKASVGERTALKKKGLPPHKLDENLMVELFHMVERHLNRVWWSPNEALS RGRVIGRP
ABM69266.1	**BmrI**	54	>gi|123187377|gb|ABM69266.1| BmrI [Bacillus megaterium] MNYFSLHPNVYATGRPKGLINMLESVWISNQKPGDGTMYLISGFANYNGGIRFYETFTEHINHGGKVIAI LGGSTSQRLSSKQVVAELVSRGVDVYIINRKRLLHAKLYGSSSNSGESLVVSSGNFTGPGMSQNVEASLL LDNNTTSSMGFSWNGMVNSMLDQKWQIHNLSNSNPTSPSWNLLYDERTTNLTLDDTQKVTLILTLGHADT ARIQAAPKSKAGEGSQYFWLSKDSYDFFPPLTIRNKRGTKATYSCLINMNYLDIKYIDSECRVTFEAENN FDFRLGTGKLRYTNVAASDDIAAITRVGDSDYELRIIKKGSSNYDALDSAAVNFIGNRGKRYGYIPNDEF GRIIGAKF
CAC12783.1	**BfiI**	55	>gi|10798463|emb|CAC12783.1| restriction endonuclease BfiI [Bacillus firmus] MNFFSLHPNVYATGRPKGLIGMLENVWVSNHTPGEGTLYLISGFSNYNGGVRFYETFTEHINQGGRVIAI LGGSTSQRLSSRQVVEELLNRGVEVHIINRKRILHAKLYGTSNNLGESLVVSSGNFTGPGMSQNIEASLL LDNNTTQSMGFSWNDMISEMLNQNWHIHNMTNATDASPGWNLLYDERTTNLTLDETERVTLIVTLGHADT ARIQAAPGTTAGQGTQYFWLSKDSYDFFPPLTIRNRRGTKATYSSLINMNYIDINYTDTQCRVTFEAENN FDFRLGTGKLRYTGVAKSNDIAAITRVGDSDYELRIIKQGTPEHSQLDPYAVSFIGNRGKRFGYISNEEF GRIIGVTF
P05725.1	**I-CreI**	1	>gi|140470|sp|P05725.1|DNE1_CHLRE RecName: Full=DNA endonuclease I- CreI; AltName: Full=23S rRNA intron protein MNTKYNKEFLLYLAGFVDGDGSIIAQIKPNQSYKFKHQLSLAFQVTQKTQRRWFLDKLVDEIGVGYVRDR GSVSDYILSEIKPLHNFLTQLQPFLKLKQKQANLVLKIIWRLPSAKESPDKFLEVCTWVDQIAALNDSKT RKTTSETVRAVLDSLSEKKKSSP
Q9UQ84.2	**hExoI (EXO1_HUMAN)**	56	>gi|85700954|sp|Q9UQ84.2|EXO1_HUMAN RecName: Full=Exonuclease 1; Short=hExo1; AltName: Full=Exonuclease I; Short=hExoI MGIQGLLQFIKEASEPIHVRKYKGQVVAVDTYCWLHKGAIACAEKLAKGEPTDRYVGFCMKFVNMLLSHG IKPILVFDGCTLPSKKEVERSRRERRQANLLKGKQLLREGKVSEARECFTRSINITHAMAHKVIKAARSQ GVDCLVAPYEADAQLAYLNKAGIVQAIITEDSDLLAFGCKKVILKMDQFGNGLEIDQARLGMCRQLGDVF TEEKFRYMCILSGCDYLSSLRGIGLAKACKVLRLANNPDIVKVIKKIGHYLKMNITVPEDYINGFIRANN

			TFLYQLVFDPIKRKLIPLNAYEDDVDPETLSYAGQYVDDSIALQIALGNKDINTFEQIDDYNPDTAMPAH SRSHSWDDKTCQKSANVSSIWHRNYSPRPESGTVSDAPQLKENPSTVGVERVISTKGLNLPRKSSIVKRP RSAELSEDDLLSQYSLSFTKKTKKNSSEGNKSLSFSEVFVPDLVNGPTNKKSVSTPPRTRNKFATFLQRK NEESGAVVVPGTRSRFFCSSDSTDCVSNKVSIQPLDETAVTDKENNLHESEYGDQEGKRLVDTDVARNSS DDIPNNHIPGDHIPDKATVFTDEESYSFESSKFTRTISPPTLGTLRSCFSWSGGLGDFSRTPSPSPSTAL QQFRRKSDSPTSLPENNMSDVSQLKSEESSDDESHPLREEACSSQSQESGEFSLQSSNASKLSQCSSKDS DSEESDCNIKLLDSQSDQTSKLRLSHFSKKDTPLRNKVPGLYKSSSADSLSTTKIKPLGPARASGLSKKP ASIQKRKHHNAENKPGLQIKLNELWKNFGFKKDSEKLPPCKKPLSPVRDNIQLTPEAEEDIFNKPECGRV QRAIFQ
P39875.2	**Yeast ExoI (EXO1_YEAST)**	57	>gi|1706421|sp|P39875.2|EXO1_YEAST RecName: Full=Exodeoxyribonuclease 1; AltName: Full=Exodeoxyribonuclease I; Short=EXO I; Short=Exonuclease I; AltName: Full=Protein DHS1 MGIQGLLPQLKPIQNPVSLRRYEGEVLAIDGYAWLHRAACSCAYELAMGKPTDKYLQFFIKRFSLLKTFK VEPYLVFDGDAIPVKKSTESKRRDKRKENKAIAERLWACGEKKNAMDYFQKCVDITPEMAKCIICYCKLN GIRYIVAPFEADSQMVYLEQKNIVQGIISEDSDLLVFGCRRLITKLNDYGECLEICRDNFIKLPKKFPLG SLTNEEIITMVCLSGCDYTNGIPKVGLITAMKLVRRFNTIERIILSIQREGKLMIPDTYINEYEAAVLAF QFQRVFCPIRKKIVSLNEIPLYLKDTESKRKRLYACIGFVIHRETQKKQIVHFDDDIDHHLHLKIAQGDL NPYDFHQPLANREHKLQLASKSNIEFGKTNTTNSEAKVKPIESFFQKMTKLDHNPKVANNIHSLRQAEDK LTMAIKRRKLSNANVVQETLKDTRSKFFNKPSMTVVENFKEKGDSIQDFKEDTNSQSLEEPVSESQLSTQ IPSSFITTNLEDDDNLSEEVSEVVSDIEEDRKNSEGKTIGNEIYNTDDDGDGDTSEDYSETAESRVPTSS TTSFPGSSQRSISGCTKVLQKFRYSSSFSGVNANRQPLFPRHVNQKSRGMVYVNQNRDDDCDDNDGKNQI TQRPSLRKSLIGARSQRIVIDMKSVDERKSFNSSPILHEESKKRDIETTKSSQARPAVRSISLLSQFVYK GK
BAJ43803.1	**E.coli ExoI**	58	>gi|315136644|dbj|BAJ43803.1| exonuclease I [Escherichia coli DH1] MMNDGKQQSTFLFHDYETFGTHPALDRPAQFAAIRTDSEFNVIGEPEVFYCKPADDYLPQPGAVLITGIT PQEARAKGENEAAFAARIHSLFTVPKTCILGYNNVRFDDEVTRNIFYRNFYDPYAWSWQHDNSRWDLLDV MRACYALRPEGINWPENDDGLPSFRLEHLTKANGIEHSNAHDAMADVYATIAMAKLVKTRQPRLFDYLFT HRNKHKLMALIDVPQMKPLVHVSGMFGAWRGNTSWVAPLAWHPENRNAVIMVDLAGDISPLLELDSDTLR ERLYTAKTDLGDNAAVPVKLVHINKCPVLAQANTLRPEDADRLGINRQHCLDNLKILRENPQVREKVVAI FAEAEPFTPSDNVDAQLYNGFFSDADRAAMKIVLETEPRNLPALDITFVDKRIEKLLFNYRARNFPGTLD YAEQQRWLEHRRQVFTPEFLQGYADELQMLVQQYADDKEKVALLKALWQYAEEIV
Q9BQ50.1	**Human TREX2**	59	>gi|47606206|sp|Q9BQ50.1|TREX2_HUMAN RecName: Full=Three prime repair exonuclease 2; AltName: Full=3'-5' exonuclease TREX2 MGRAGSPLPRSSWPRMDDCGSRSRCSPTLCSSLRTCYPRGNITMSEAPRAETFVFLDLEATGLPSVEPEI AELSLFAVHRSSLENPEHDESGALVLPRVLDKLTLCMCPERPFTAKASEITGLSSEGLARCRKAGFDGAV VRTLQAFLSRQAGPICLVAHNGFDYDFPLLCAELRRLGARLPRDTVCLDTLPALRGLDRAHSHGTRARGR QGYSLGSLFHRYFRAEPSAAHSAEGDVHTLLLIFLHRAAELLAWADEQARGWAHIEPMYLPPDDPSLEA
Q91XB0.2	**Mouse TREX1**	60	>gi|47606196|sp|Q91XB0.2|TREX1_MOUSE RecName: Full=Three prime repair exonuclease 1; AltName: Full=3'-5' exonuclease TREX1 MGSQTLPHGHMQTLIFLDLEATGLPSSRPEVTELCLLAVHRRALENTSISQGHPPPVPRPPRVVDKLSLC IAPGKACSPGASEITGLSKAELEVQGRQRFDDNLAILLRAFLQRQPQPCCLVAHNGDRYDFPLLQTELAR LSTPSPLDGTFCVDSIAALKALEQASSPSGNGSRKSYSLGSIYTRLYWQAPTDSHTAEGDVLTLLSICQW KPQALLQWVDEHARPFSTVKPMYGTPATTGTTNLRPHAATATTPLATANGSPSNGRSRRPKSPPPEKVPE APSQEGLLAPLSLLTLLTLAIATLYGLFLASPGQ
Q9NSU2.1	**Human TREX1**	61	>gi|47606216|sp|Q9NSU2.1|TREX1_HUMAN RecName: Full=Three prime repair exonuclease 1; AltName: Full=3'-5' exonuclease TREX1; AltName: Full=DNase III MGPGARRQGRIVQGRPEMCFCPPPTPLPPLRILTLGTHTPTPCSSPGSAAGTYPTMGSQALPPGPMQTLI FFDMEATGLPFSQPKVTELCLLAVHRCALESPPTSQGPPPTVPPPPRVVDKLSLCVAPGKACSPAASEIT GLSTAVLAAHGRQCFDDNLANLLLAFLRRQPQPWCLVAHNGDRYDFPLLQAELAMLGLTSALDGAFCVDS ITALKALERASSPSEHGPRKSYSLGSIYTRLYGQSPPDSHTAEGDVLALLSICQWRPQALLRWVDAHARP FGTIRPMYGVTASARTKPRPSAVTTTAHLATTRNTSPSLGESRGTKDLPPVKDPGALSREGLLAPLGLLA ILTLAVATLYGLSLATPGE
Q9BG99.1	**Bovine TREX1**	62	>gi|47606205|sp|Q9BG99.1|TREX1_BOVIN RecName: Full=Three prime repair exonuclease 1; AltName: Full=3'-5' exonuclease TREX1 MGSRALPPGPVQTLIFLDLEATGLPFSQPKITELCLLAVHRYALEGLSAPQGPSPTAPVPPRVLDKLSLC VAPGKVCSPAASEITGLSTAVLAAHGRRAFDADLVNLIRTFLQRQPQPWCLVAHNGDRYDFPLLRAELAL LGLASALDDAFCVDSIAALKALEPTGSSSEHGPRKSYSLGSVYTRLYGQAPPDSHTAEGDVLALLSVCQW RPRALLRWVDAHAKPFSTVKPMYVITTSTGTNPRPSAVTATVPLARASDTGPNLRGDRSPKPAPSPKMCP GAPPGEGLLAPLGLLAFLTLAVAMLYGLSLAMPGQ
AAH91242.1	**Rat TREX1**	63	>gi|60688197|gb|AAH91242.1| Trex1 protein [Rattus norvegicus] MGSQALPHGHMQTLIFLDLEATGLPYSQPKITELCLLAVHRHALENSSMSEGQPPPVPKPPRVVDKLSLC IAPGKPCSSGASEITGLTTAGLEAHGRQRFNDNLATLLQVFLQRQPQPCCLVAHNGDRYDFPLLQAELAS LSVISPLDGTFCVDSIAALKTLEQASSPSEHGPRKSYSLGSIYTRLYGQAPTDSHTAEGDVLALLSICQW KPQALLQWVDKHARPFSTIKPMYGMAATTGTASPRLCAATTSSPLATANLSPSNGRSRGKRPTSPPPENV PEAPSREGLLAPLGLLTFLTLAIAVLYGIFLASPGQ
AAH63664.1	**Human DNA2**	64	>gi|39793966|gb|AAH63664.1| **DNA2 protein** [Homo sapiens] FAIPASRMEQLNELELLMEKSFWEEAELPAELFQKKVVASFPRTVLSTGMDNRYLVLAVNTVQNKEGNCE KRLVITASQSLENKELCILRNDWCSVPVEPGDIIHLEGDCTSDTWIIDKDFGYLILYPDMLISGTSIASS IRCMRRAVLSETFRSSDPATRQMLIGTVLHEVFQKAINNSFAPEKLQELAFQTIQEIRHLKEMYRLNLSQ DEIKQEVEDYLPSFCKWAGDFMHKNTSTDFPQMQLSLPSDNSKDNSTCNIEVVKPMDIEESIWSPRFGLK GKIDVTVGVKIHRGYKTKYKIMPLELKTGKESNSIEHRSQVVLYTLLSQERRADPEAGLLLYLKTGQMYP VPANHLDKRELLKLRNQMAFSLFHRISKSATRQKTQLASLPQIIEEEKTCKYCSQIGNCALYSRAVEQQM DCSSVPIVMLPKIEEETQHLKQTHLEYFSLWCLMLTLESQSKDNKKNHQNIWLMPASEMEKSGSCIGNLI RMEHVKIVCDGQYLHNFQCKHGAIPVTNLMAGDRVIVSGEERSLFALSRGYVKEINMTTVTCLLDRNLSV LPESTLFRLDQEEKNCDIDTPLGNLSKLMENTFVSKKLRDLIIDFREPQFISYLSSVLPHDAKDTVACIL KGLNKPQRQAMKKVLLSKDYTLIVGMPGTGKTTTICTLVPAPEQVEKGGVSNVTEAKLIVFLTSIFVKAG

			CSPSDIGIIAPYRQQLKIINDLLARSIGMVEVNTVDKYQGRDKSIVLVSFVRSNKDGTVGELLKDWRRLN VAITRAKHKLILLGCVPSLNCYPPLEKLLNHLNSEKLIIDLPSREHESLCHILGDFQRE
P38859.1	**Yeast DNA2 (DNA2_YEAST)**	65	>gi|731738|sp|P38859.1|DNA2_**YEAST** RecName: Full=DNA replication ATP-dependent helicase **DNA2** MPGTPQKNKRSASISVSPAKKTEEKEIIQNDSKAILSKQTKRKKKYAFAPINNLNGKNTKVSNASVLKSI AVSQVRNTSRTKDINKAVSKSVKQLPNSQVKPKREMSNLSRHHDFTQDEDGPMEEVIWKYSPLQRDMSDK TTSAAEYSDDYEDVQNPSSTPIVPNRLKTVLSFTNIQVPNADVNQLIQENGNEQVRPKPAEISTRESLRN IDDILDDIEGDLTIKPTITKFSDLPSSPIKAPNVEKKAEVNAEEVDKMDSTGDSNDGDDSLIDILTQKYV EKRKSESQITIQGNTNQKSGAQESCGKNDNTKSRGEIEDHENVDNQAKTGNAFYENEEDSNCQRIKKNEK IEYNSSDEFSDDSLIELLNETQTQVEPNTIEQDLDKVEKMVSDDLRIATDSTLSAYALRAKSGAPRDGVV RLVIVSLRSVELPKIGTQKILECIDGKGEQSSVVVRHPWVYLEFEVGDVIHIIEGKNIENKRLLSDDKNP KTQLANDNLLVLNPDVLFSATSVGSSVGCLRRSILQMQFQDPRGEPSLVMTLGNIVHELLQDSIKYKLSH NKISMEIIIQKLDSLLETYSFSIIICNEEIQYVKELVMKEHAENILYFVNKFVSKSNYGCYTSISGTRRT QPISISNVIDIEENIWSPIYGLKGFLDATVEANVENNKKHIVPLEVKTGKSRSVSYEVQGLIYTLLLNDR YEIPIEFFLLYFTRDKNMTKFPSVLHSIKHILMSRNRMSMNFKHQLQEVFGQAQSRFELPPLLRDSSCDS CFIKESCMVLNKLLEDGTPEESGLVEGEFEILTNHLSQNLANYKEFFTKYNDLITKEESSITCVNKELFL LDGSTRESRSGRCLSGLVVSEVVEHEKTEGAYIYCFSRRRNDNNSQSMLSSQIAANDFVIISDEEGHFCL CQGRVQFINPAKIGISVKRKLLNNRLLDKEKGVTTIQSVVESELEQSSLIATQNLVTYRIDKNDIQQSLS LARFNLLSLFLPAVSPGVDIVDERSKLCRKTKRSDGGNEILRSLLVDNRAPKFRDANDDPVIPYKLSKDT TLNLNQKEAIDKVMRAEDYALILGMPGTGKTTVIAEIIKILVSEGKRVLLTSYTHSAVDNILIKLRNTNI SIMRLGMKHKVHPDTQKYVPNYASVKSYNDYLSKINSTSVVATTCLGINDILFTLNEKDFDYVILDEASQ ISMPVALGPLRYGNRFIMVGDHYQLPPLVKNDAARLGGLEESLFKTFCEKHPESVAELTLQYRMCGDIVT LSNFLIYDNKLKCGNNEVFAQSLELPMPEALSRYRNESANSKQWLEDILEPTRKVVFLNYDNCPDIIEQS EKDNITNHGEAELTLQCVEGMLLSGVPCEDIGVMTLYRAQLRLLKKIFNKNVYDGLEILTADQFQGRDKK CIIISMVRRNSQLNGGALLKELRRVNVAMTRAKSKLIIIGSKSTIGSVPEIKSFVNLLEERNWVYTMCKD ALYKYKFPDRSNAIDEARKGCGKRTGAKPITSKSKFVSDKPIIKEILQEYES
AAA45863.1	**VP16**	66	>gi|330318|gb|AAA45863.1| **VP16** [Human herpesvirus 2] MDLLVDDLFADRDGVSPPPPRPAGGPKNTPAAPPLYATGRLSQAQLMPSPPMPVPPAALFNRLLDDLGFS AGPALCTMLDTWNEDLFSGFPTNADMYRECKFLSTLPSDVIDWGDAHVPERSPIDIRAHGDVAFPTLPAT RDELPSYYEAMAQFFRGELRAREESYRTVLANFCSALYRYLRASVRQLHRQAHMRGRNRDLREMLRTTIA DRYYRETARLARVLFLHLYLFLSREILWAAYAEQMMRPDLFDGLCCDLESWRQLACLFQPLMFINGSLTV RGVPVEARRLRELNHIREHLNLPLVRSAAAEEPGAPLTTPPVLQGNQARSSGYFMLLIRAKLDSYSSVAT SEGESVMREHAYSRGRTRNNYGSTIEGLLDLPDDDDAPAEAGLVAPRMSFLSAGQRPRRLSTTAPITDVS LGDELRLDGEEVDMTPADALDDFDLEMLGDVESPSPGMTHDPVSYGALDVDDFEFEQMFTDAMGIDDFGG
ACM07430.1	**Colicin E9**	366	>gi|221185856|gb|ACM07430.1| colicin E9 [Escherichia coli] MSGGDGRGHNTGAHSTSGNINGGPTGIGVSGGASDGSGWSSENNPWGGGSGSGIHWGGGSGRGNGGGNGN SGGGSGTGGNLSAVAAPVAFGFPALSTPGAGGLAVSISASELSAAIAGIIAKLKKVNLKFTPFGVVLSSL IPSEIAKDDPNMMSKIVTSLPADDITESPVSSLPLDKATVNVNVRVVDDVKDERQNISVVSGVPMSVPVV DAKPTERPGVFTASIPGAPVLNISVNDSTPAVQTLSPGVTNNTDKDVRPAGFTQGGNTRDAVIRFPKDSG HNAVYVSVSDVLSPDQVKQRQDEENRRQQEWDATHPVEAAERNYERARAELNQANEDVARNQERQAKAVQ VYNSRKSELDAANKTLADAIAEIKQFNRFAHDPMAGGHRMWQMAGLKAQRAQTDVNNKQAAFDAAAKEKS DADAALSAAQERRKQKENKEKDAKDKLDKESKRNKPGKATGKGKPVGDKWLDDAGKDSGAPIPDRIADKL RDKEFKSFDDFRKAVWEEVSKDPELSKNLNPSNKSSVSKGYSPFTPKNQQVGGRKVYELHHDKPISQGGE VYDMDNIRVTTPKRHIDIHRGK
NP_775816.1	**APFL**	367	>gi|135233|sp|P14870.1|T2F1_PLAOK RecName: Full=Type-2 restriction enzyme FokI; Short=R.FokI; AltName: Full=Endonuclease FokI; AltName: Full=Type II restriction enzyme FokI; AltName: Full=Type IIS restriction enzyme FokI MFLSMVSKIRTFGWVQNPGKFENLKRVVQVFDRNSKVHNEVKNIKIPTLVKESKIQKELVAIMNQHDLIY TYKELVGTGTSIRSEAPCDAIIQATIADQGNKKGYIDNWSSDGFLRWAHALGFIEYINKSDSFVITDVGL AYSKSADGSAIEKEILIEAISSYPPAIRILTLLEDGQHLTKFDLGKNLGFSGESGFTSLPEGILLDTLAN AMPKDKGEIRNNWEGSSDKYARMIGGWLDKLGLVKQGKKEFIIPTLGKPDNKEFISHAFKITGEGLKVLR RAKGSTKFTRVPKRVYWEMLATNLTDKEYVRTRRALILEILIKAGSLKIEQIQDNLKKLGFDEVIETIEN DIKGLINTGIFIEIKGRFYQLKDHILQFVIPNRGVTKQLVKSELEEKKSELRHKLKYVPHEYIELIEIAR NSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQ RYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGE MIKAGTLTLEEVRRKFNNGEINF
P14870.1	**FokI**	368	>gi|221185857|gb|ACM07431.1| colicin E9 immunity protein [Escherichia coli] MELKHSISDYTEAEFLQLVTTICNADTSSEEELVKLVTHFEEMTEHPSGSDLIYYPKEGDDDSPSGIVNT VKQWRAANGKSGFKQG

Table 2: List of catalytic / enhancer domains for compact TALENs or enhanced compact TALENs.

In another preferred embodiment according to the method of the present invention, the peptidic linker that can link said catalytic domain to the core TALE scaffold according to the method of the

present invention can be selected from the group consisting of NFS1, NFS2, CFS1, RM2, BQY, QGPSG, LGPDGRKA, 1a8h_1, 1dnpA_1, 1d8cA_2, 1ckqA_3, 1sbp_1, 1ev7A_1, 1alo_3, 1amf_1, 1adjA_3, 1fcdC_1, 1al3_2, 1g3p_1, 1acc_3, 1ahjB_1, 1acc_1, 1af7_1, 1heiA_1, 1bia_2, 1igtB_1, 1nfkA_1, 1au7A_1, 1bpoB_1, 1b0pA_2, 1c05A_2, 1gcb_1, 1bt3A_1, 1b3oB_2, 16vpA_6, 1dhx_1, 1b8aA_1and 1qu6A_1, as listed in Table 3 (SEQ ID NO: 67 to SEQ ID NO: 104 and SEQ ID NO: 372 to SEQ ID NO: 415). In a more preferred embodiment, the peptidic linker that can link said catalytic domain to the core TALE scaffold according to the method of the present invention can be selected from the group consisting of NFS1 (SEQ ID NO: 98), NFS2 (SEQ ID NO: 99) andCFS1 (SEQ ID NO: 100). In the scope of the present invention is also encompassed the case where a peptidic linker is not needed to fuse a catalytical domain to the TALE scaffold in order to obtain a cTALEN according to the present invention.

Name (PDB)	Amino Acids	Size	Length	Sequence	SEQ ID NO
1a8h_1	285 – 287	3	6,636	NVG	67
1dnpA_1	130 – 133	4	7,422	DSVI	68
1d8cA_2	260 – 263	4	8,782	IVEA	69
1ckqA_3	169 – 172	4	9,91	LEGS	70
1sbp_1	93 – 96	4	10,718	YTST	71
1ev7A_1	169 – 173	5	11,461	LQENL	72
1alo_3	360 – 364	5	12,051	VGRQP	73
1amf_1	81 – 85	5	13,501	LGNSL	74
1adjA_3	323 – 328	6	14,835	LPEEKG	75
1fcdC_1	76 – 81	6	14,887	QTYQPA	76
1al3_2	265 – 270	6	15,485	FSHSTT	77
1g3p_1	99 – 105	7	17,903	GYTYINP	78
1acc_3	216 – 222	7	19,729	LTKYKSS	79
1ahjB_1	106 – 113	8	17,435	SRPSESEG	80
1acc_1	154 – 161	8	18,776	PELKQKSS	81
1af7_1	89 – 96	8	22,502	LTTNLTAF	82
1heiA_1	322 – 330	9	13,534	TATPPGSVT	83
1bia_2	268 – 276	9	16,089	LDNFINRPV	84
1igtB_1	111 – 119	9	19,737	VSSAKTTAP	85
1nfkA_1	239 – 248	10	13,228	DSKAPNASNL	86
1au7A_1	103 – 112	10	20,486	KRRTTISIAA	87
1bpoB_1	138 – 148	11	21,645	PVKMFDRHSSL	88
1b0pA_2	625 – 635	11	26,462	APAETKAEPMT	89
1c05A_2	135 – 148	14	23,819	YTRLPERSELPAEI	90
1gcb_1	57 – 70	14	27,39	VSTDSTPVTNQKSS	91
1bt3A_1	38 – 51	14	28,818	YKLPAVTTMKVRPA	92
1b3oB_2	222 – 236	15	20,054	IARTDLKKNRDYPLA	93

16vpA_6	312 – 332	21	23,713	TEEPGAPLTTPPTLHGNQARA	94
1dhx_1	81 – 101	21	42,703	ARFTLAVGDNRVLDMASTYFD	95
1b8aA_1	95 – 120	26	31,305	IVVLNRAETPLPLDPTGKVKAELDTR	96
1qu6A_1	79 – 106	28	51,301	ILNKEKKAVSPLLLTTTNSSEGLSMGNY	97
NFS1	-	20	-	GSDITKSKISEKMKGQGPSG	98
NFS2	-	23	-	GSDITKSKISEKMKGLGPDGRKA	99
CFS1	-	10	-	SLTKSKISGS	100
RM2	-	32	-	AAGGSALTAGALSLTAGALSLTAGALSGGGGS	101
BQY	-	25	-	AAGASSVSASGHIAPLSLPSSPPSVGS	102
QGPSG	-	5	-	QGPSG	103
LGPDGRKA	-	8	-	LGPDGRKA	104

TAL1	-	15	-	SGGSGSNVGSGSGSG	372
TAL2	-	20	-	SGGSGSLTTNLTAFSGSGSG	373
TAL3	-	22	-	SGGSGSKRRTTISIAASGSGSG	374
TAL4	-	17	-	SGGSGSVGRQPSGSGSG	375
TAL5	-	26	-	SGGSGSYTRLPERSELPAEISGSGSG	376
TAL6	-	38	-	SGGSGSIVVLNRAETPLPLDPTGKVKAELDTRSGSGSG	377
TAL7	-	21	-	SGGSGSTATPPGSVTSGSGSG	378
TAL8	-	21	-	SGGSGSLDNFINRPVSGSGSG	379
TAL9	-	21	-	SGGSGSVSSAKTTAPSGSGSG	380
TAL10	-	22	-	SGGSGSDSKAPNASNLSGSGSG	381
TAL11	-	23	-	SGGSGSPVKMFDRHSSLSGSGSG	382
TAL12	-	23	-	SGGSGSAPAETKAEPMTSGSGSG	383
TAL13	-	26	-	SGGSGSVSTDSTPVTNQKSSSGSGSG	384
TAL14	-	16	-	SGGSGSDSVISGSGSG	385
TAL15	-	33	-	SGGSGSARFTLAVGDNRVLDMASTYFDSGSGSG	386
TAL16	-	17	-	SGGSGSLQENLSGSGSG	387
TAL17	-	19	-	SGGSGSGYTYINPSGSGSG	388
TAL18	-	26	-	SGGSGSYKLPAVTTMKVRPASGSGSG	389
TAL19	-	16	-	SGGSGSLEGSSGSGSG	390
TAL20	-	16	-	SGGSGSIVEASGSGSG	391
TAL21	-	18	-	SGGSGSQTYQPASGSGSG	392
TAL22	-	27	-	SGGSGSIARTDLKKNRDYPLASGSGSG	393
TAL23	-	18	-	SGGSGSLPEEKGSGSGSG	394
TAL24	-	16	-	SGGSGSYTSTSGSGSG	395
TAL25	-	20	-	SGGSGSSRPSESEGSGSGSG	396
TAL26	-	17	-	SGGSGSLGNSLSGSGSG	397
TAL27	-	19	-	SGGSGSLTKYKSSSGSGSG	398
TAL28	-	33	-	SGGSGSTEEPGAPLTTPPTLHGNQARASGSGSG	399
TAL29	-	18	-	SGGSGSFSHSTTSGSGSG	400
TAL30	-	20	-	SGGSGSPELKQKSSSGSGSG	401
TAL31	-	40	-	SGGSGSILNKEKKAVSPLLLTTTNSSEGLSMGNYSGSGSG	402

TAL32	-	31	-	ELAEFHARYADLLLRDLRERPVSLVRGPDSG	403
TAL33	-	31	-	ELAEFHARPDPLLLRDLRERPVSLVRGLGSG	404
TAL34	-	26	-	ELAEFHARYADLLLRDLRERSGSGSG	405
TAL35	-	31	-	DIFDYYAGVAEVMLGHIAGRPATRKRWPNSG	406
TAL36	-	31	-	DIFDYYAGPDPVMLGHIAGRPATRKRWLGSG	407
TAL37	-	26	-	DIFDYYAGVAEVMLGHIAGRSGSGSG	408
Linker A		37		SIVAQLSRPDPALVSFQKLKLACLGGRPALDAVKKGL	409
Linker B		37		SIVAQLSRPDPAAVSAQKAKAACLGGRPALDAVKKGL	410
Linker C		37		SIVAQLSRPDPAVVTFHKLKLACLGGRPALDAVKKGL	411
Linker D		44		SIVAQLSRPDPAQSLAQELSLNESQIKIACLGGRPALDAVKKGL	412
Linker E		40		SIVAQLSRPDPALQLPPLERLTLDACLGGRPALDAVKKGL	413
Linker F		38		SIVAQLSRPDPAIHKKFSSIQMACLGGRPALDAVKKGL	414
Linker G		40		SIVAQLSRPDPAAAAATNDHAVAAACLGGRPALDAVKKGL	415

Table 3: List of peptidic linkers that can be used in compact TALENs or enhanced compact TALENs.

Depending from its structural composition [type of core TALE scaffold, type of catalytic domain(s) with associated enzymatic activities and eventually type of linker(s)], a compact TALEN according to the present invention can comprise different levels of separate enzymatic activities able to differently process DNA, resulting in a global DNA processing efficiency for said compact TALEN, each one of said different enzymatic activities having their own DNA processing efficiency.

In another preferred embodiment, the method according to the present invention further comprises the steps of:

(i) Engineering at least one enhancer domain;

(ii) Optionally determining or engineering one peptide linker to fuse said enhancer domain to one part of said compact TALEN entity;

thereby obtaining a compact TALEN entity with enhanced DNA processing efficiency nearby a single double-stranded DNA target sequence of interest, i.e an enhanced compact TALEN.

In other words, according to the method of the present invention said unique compact TALEN monomer further comprises:

(i) At least one enhancer domain;

(ii) Optionally one peptide linker to fuse said enhancer domain to one part of said unique compact TALEN monomer active entity.

In another more preferred embodiment, said enhancer domain is fused to the N-terminus of the core TALE scaffold part of said compact TALEN entity. In another more preferred embodiment, said

enhancer domain is fused to C-terminus of the core TALE scaffold part of said compact TALEN entity. In another more preferred embodiment, said enhancer domain is fused to the catalytic domain part of said compact TALEN entity. In another more preferred embodiment, said enhancer domain is fused between the N-terminus of the core TALE scaffold part and the catalytic part of said compact TALEN entity. In another more preferred embodiment, said enhancer domain is fused between the C-terminus of the core TALE scaffold part and the catalytic part of said compact TALEN entity. In the scope of the present invention, it can be envisioned to insert said catalytic domain and / or enhancer domain between two parts of the engineered core TALE scaffold according to the invention, each part comprising one set of RVDs. In this last case, the number of RVDs for each engineered core TALE scaffold can be the same or not. In other words, it can be envisioned to split said core TALE scaffold of the present invention to insert one catalytic domain and / or one enhancer domain between the resulting two parts of said engineered core TALE scaffold.

In another preferred embodiment, said enhancer domain is catalytically active or not, providing functional and/or structural support to said compact TALEN entity. In a more preferred embodiment, said enhancer domain consists of a protein domain selected from the group consisting of MmeI, Colicin-E7 (CEA7_ECOLX), Colicin-E9, APFL, EndA, Endo I (END1_ECOLI), Human Endo G (NUCG_HUMAN), Bovine Endo G (NUCG_BOVIN), R.HinP1I, I-BasI, I-BmoI, I-HmuI, I-TevI, I-TevII, I-TevIII, I-TwoI, R.MspI, R.MvaI, NucA, NucM, Vvn, Vvn_CLS, Staphylococcal nuclease (NUC_STAAU), Staphylococcal nuclease (NUC_STAHY), Micrococcal nuclease (NUC_SHIFL), Endonuclease yncB, Endodeoxyribonuclease I (ENRN_BPT7), Metnase, Nb.BsrDI, BsrDI A, Nt.BspD6I (R.BspD6I large subunit), ss.BspD6I (R.BspD6I small subunit), R.PleI, MlyI, AlwI, Mva1269I, BsrI, BsmI, Nb.BtsCI, Nt.BtsCI, R1.BtsI, R2.BtsI, BbvCI subunit 1, BbvCI subunit 2, Bpu10I alpha subunit, Bpu10I beta subunit, BmrI, BfiI, I-CreI, hExoI (EXO1_HUMAN), Yeast ExoI (EXO1_YEAST), E.coli ExoI, Human TREX2, Mouse TREX1, Human TREX1, Bovine TREX1, Rat TREX1, Human DNA2, Yeast DNA2 (DNA2_YEAST) and VP16, as listed in Table 2 (SEQ ID NO: 10 to SEQ ID NO: 66 and SEQ ID NO: 1, 366 & 367, a functional mutant, a variant or a derivative thereof. In another more preferred embodiment, said enhancer domain consists of a catalytically active derivative of the protein domains listed above and in Table 2, providing functional and/or structural support to said compact TALEN entity. In another preferred embodiment, said enhancer domain consists of a catalytically inactive derivative of the protein domains listed above and in Table 2, providing structural support to said compact TALEN entity. In another preferred embodiment, said enhancer domain is selected from the group consisting of I-TevI (SEQ ID NO: 20), ColE7 (SEQ ID NO: 11) and NucA (SEQ ID NO: 26).

In a more preferred embodiment, said enhanced compact TALEN according to the method of the present invention can comprise a second enhancer domain. In this embodiment, said second enhancer domain can have the same characteristics than the first enhancer domain. In a more preferred embodiment, said second enhancer domain provides structural support to enhanced compact TALEN entity. In another more preferred embodiment, said second enhancer domain provides functional support to enhanced compact TALEN entity. In a more preferred embodiment, said second enhancer domain provides structural and functional support to the enhanced compact TALEN entity. In a more preferred embodiment, said enhanced compact TALEN entity comprises one catalytic domain and one enhancer domain. In another more preferred embodiment said enhanced compact TALEN entity comprises one catalytic domain and two enhancer domains. In another more preferred embodiment said enhanced compact TALEN entity comprises two catalytic domains and one enhancer domains. In another more preferred embodiment said enhanced compact TALEN entity comprises two catalytic domains and two enhancer domains.

In a more preferred embodiment, said second enhancer domain consists of a protein domain derived from a protein selected from the group consisting of MmeI, Colicin-E7 (CEA7_ECOLX), Colicin-E9, APFL, EndA, Endo I (END1_ECOLI), Human Endo G (NUCG_HUMAN), Bovine Endo G (NUCG_BOVIN), R.HinP1I, I-BasI, I-BmoI, I-HmuI, I-TevI, I-TevII, I-TevIII, I-TwoI, R.MspI, R.MvaI, NucA, NucM, Vvn, Vvn_CLS, Staphylococcal nuclease (NUC_STAAU), Staphylococcal nuclease (NUC_STAHY), Micrococcal nuclease (NUC_SHIFL), Endonuclease yncB, Endodeoxyribonuclease I (ENRN_BPT7), Metnase, Nb.BsrDI, BsrDI A, Nt.BspD6I (R.BspD6I large subunit), ss.BspD6I (R.BspD6I small subunit), R.PleI, MlyI, AlwI, Mva1269I, BsrI, BsmI, Nb.BtsCI, Nt.BtsCI, R1.BtsI, R2.BtsI, BbvCI subunit 1, BbvCI subunit 2, Bpu10I alpha subunit, Bpu10I beta subunit, BmrI, BfiI, I-CreI, hExoI (EXO1_HUMAN), Yeast ExoI (EXO1_YEAST), E.coli ExoI, Human TREX2, Mouse TREX1, Human TREX1, Bovine TREX1, Rat TREX1, Human DNA2, Yeast DNA2 (DNA2_YEAST) and VP16, as listed in Table 2 (SEQ ID NO: 10 to SEQ ID NO: 66 and SEQ ID NO: 1, 366 & 367, a functional mutant, a variant or a derivative thereof. In another more preferred embodiment, said second enhancer domain consists of a catalytically active derivative of the protein domains listed above and in Table 2, providing functional and/or structural support to said enhanced compact TALEN entity. In another preferred embodiment, said second enhancer domain consists of a catalytically inactive derivative of the protein domains listed above and in Table 2, providing structural support to said enhanced compact TALEN entity.

In another more preferred embodiment, any combinations of catalytic and/or enhancer domains listed above, as non-limiting examples, can be envisioned to be fused to said core TALE scaffold providing structural and/or functional support to said compact TALEN entity. More preferably, combinations of catalytic domains selected from the group of TevI (SEQ ID NO: 20), ColE7 (SEQ ID

NO: 11) and NucA (SEQ ID NO: 26) can be envisioned. Optionally, FokI (SEQ ID NO: 368) can be used in combination with another catalytic domain according to the list of Table2. Such combinations of catalytic and/or enhancer domains can be envisioned regarding the envisioned applications for using the method of the present invention.

Depending from its structural composition [type of core TALE scaffold, type of catalytic domain(s) with associated enzymatic activities, eventually type of linker(s) and type of enhancer(s) domains], an enhanced compact TALEN according to the present invention can present different levels of separate enzymatic activities able to differently process DNA, resulting in a global DNA processing efficiency for said enhanced compact TALEN, each one of said different enzymatic activities having their own DNA processing efficiency.

In this preferred embodiment, the DNA processing efficiency of the compact TALEN entity according to the method of the present invention can be enhanced by the engineering of at least one enhancer domain and one peptidic linker thereby obtaining a compact TALEN entity with enhanced DNA processing activity nearby a single double-stranded DNA target sequence of interest, i.e a enhanced compact TALEN according to the present invention.

Depending on its structural composition, the global DNA processing efficiency that is enhanced in said enhanced compact TALEN according to the present invention, can have a dominant enzymatic activity selected from the group consisting of a nuclease activity, a polymerase activity, a kinase activity, a phosphatase activity, a methylase activity, a topoisomerase activity, an integrase activity, a transposase activity or a ligase activity as non-limiting examples. In a more preferred embodiment, the global DNA processing efficiency that is enhanced in said enhanced compact TALEN according to the present invention is a combination of different enzymatic activities selected from the group consisting of a nuclease activity, a polymerase activity, a kinase activity, a phosphatase activity, a methylase activity, a topoisomerase activity, an integrase activity, a transposase activity or a ligase activity as non-limiting examples. In a more preferred embodiment, the global DNA processing efficiency that is enhanced in said enhanced compact TALEN according to the present invention is one of its different enzymatic activities selected from the group consisting of a nuclease activity, a polymerase activity, a kinase activity, a phosphatase activity, a methylase activity, a topoisomerase activity, an integrase activity, a transposase activity or a ligase activity as non-limiting examples. In this case, the global DNA processing efficiency is equivalent to one DNA processing activity amongst the enzymatic activities mentioned above. In another more preferred embodiment, said DNA

processing activity of the compact TALEN entity which is enhanced by the enhancer is a cleavase activity or a nickase activity or a combination of both a cleavase activity and a nickase activity.

Enhancement of DNA processing efficiency of a compact TALEN entity according to the present invention can be a consequence of a structural support by at least one enhancer domain. In a preferred embodiment, said structural support enhances the binding of a compact TALEN entity according to the invention for said DNA target sequence compared to the binding of a starting compact TALEN entity for the same DNA target sequence, thereby indirectly assisting the catalytic domain(s) to obtain a compact TALEN entity with enhanced DNA processing activity. In another preferred embodiment, said structural support enhances the existing catalytical activity of a compact TALEN entity for a DNA target sequence compared to the binding of a starting compact TALEN entity for the same DNA target sequence to obtain a compact TALEN entity with enhanced DNA processing activity.
In another preferred embodiment, said enhancer according to the method of the present invention both enhances the binding of the compact TALEN entity for said DNA target sequence and the catalytic activity of the catalytic domain(s) to obtain a compact TALEN entity with enhanced DNA processing activity. All these non-limiting examples lead to a compact TALEN entity with enhanced DNA processing efficiency for a DNA target sequence at a genomic locus of interest, i.e an enhanced compact TALEN according to the present invention.

Enhancement of DNA processing efficiency of a compact TALEN entity according to the present invention, compared to a starting compact TALEN entity, can also be a consequence of a fuctional support by at least one enhancer domain. In a preferred embodiment, said functional support can be the consequence of the hydrolysis of additional phosphodiester bonds. In a more preferred embodiment, said functional support can be the hydrolysis of additional phosphodiester bonds by a protein domain derived from a nuclease. In an embodiment, said functional support can be the hydrolysis of additional phosphodiester bonds by a protein domain derived from an endonuclease. In a more preferred embodiment, said functional support can be the hydrolysis of additional phosphodiester bonds by a protein domain derived from a cleavase. In another more preferred embodiment, said functional support can be the hydrolysis of additional phosphodiester bonds by a protein domain derived from a nickase. In a more preferred embodiment, said functional support can be the hydrolysis of additional phosphodiester bonds by a protein domain derived from an exonuclease.

In genome engineering experiments, the efficiency of rare-cutting endonuclease, e.g. their ability to induce a desired event (Homologous gene targeting, targeted mutagenesis, sequence removal or excision) at a locus, depends on several parameters, including the specific activity of the nuclease, probably the accessibility of the target, and the efficacy and outcome of the repair pathway(s) resulting in the desired event (homologous repair for gene targeting, NHEJ pathways for targeted mutagenesis).

Cleavage by peptidic rare cutting endonucleases usually generates cohesive ends, with 3' overhangs for LAGLIDADG meganucleases (Chevalier and Stoddard 2001) and 5' overhangs for Zinc Finger Nucleases (Smith, Bibikova et al. 2000). These ends, which result from hydrolysis of phosphodiester bonds, can be re-ligated in vivo by NHEJ in a seamless way (i.e a scarless re-ligation). The restoration of a cleavable target sequence allows for a new cleavage event by the same endonuclease, and thus, a series of futile cycles of cleavage and re-ligation events can take place. Indirect evidences have shown that even in the yeast Saccharomyces cerevisiae, such cycles could take place upon continuous cleavage by the HO endonuclease (Lee, Paques et al. 1999). In mammalian cells, several experiment have shown that perfect re-ligation of compatible cohesive ends resulting from two independent but close I-SceI-induced DSBs is an efficient process (Guirouilh-Barbat, Huck et al. 2004; Guirouilh-Barbat, Rass et al. 2007; Bennardo, Cheng et al. 2008; Bennardo, Gunn et al. 2009). Absence of the Ku DNA repair protein does not significantly affect the overall frequency of NHEJ events rejoining the ends from the two DSBs; however it very strongly enhances the contribution of imprecise NHEJ to the repair process in CHO immortalized cells and mouse ES cells (Guirouilh-Barbat, Huck et al. 2004; Guirouilh-Barbat, Rass et al. 2007; Bennardo, Cheng et al. 2008). Furthermore, the absence of Ku stimulates I-SceI-induced events such as imprecise NHEJ (Bennardo, Cheng et al. 2008), single-strand annealing (Bennardo, Cheng et al. 2008) and gene conversion (Pierce, Hu et al. 2001; Bennardo, Cheng et al. 2008) in mouse ES cells . Similar observations shave been made with cells deficient for the XRCC4 repair protein (Pierce, Hu et al. 2001; Guirouilh-Barbat, Rass et al. 2007; Bennardo, Gunn et al. 2009) (although XRCC4 deficiency affects the overal level of NHEJ in CHO cells (Guirouilh-Barbat, Rass et al. 2007)) or for DNA-PK (Pierce, Hu et al. 2001). In contrast, knock-down of CtIP has been shown to suppresses "alt-NHEJ" (a Ku- and XRCC4-independent form of NHEJ more prone to result in imprecise NHEJ), single-strand annealing and gene conversion, while not affecting the overall level of rejoining of two compatible ends generated by I-SceI (Bennardo, Cheng et al. 2008). Thus, competition between different DSB repair pathways can affect the spectrum or repair events resulting from a nuclease-induced DSB.

In addition, DSB resection is important for certain DSB pathways. Extensive DSB resection, resulting in the generation of large single stranded regions (a few hundred nucleotides at least), has been shown in yeast to initiate single strand annealing (Sugawara and Haber 1992) and strand invasion, the ATP-dependant step that initiates many homologous recombination events of DNA duplex invasion by an homologous strand that (White and Haber 1990; Sun, Treco et al. 1991) (for a review of mechanisms, see (Paques and Haber 1999)). In eukaryotic cells DSB resection depends on several proteins including BLM/Sgs1 and DNA2, EXOI, and the MRN complex (Mre11, Rad50, Nbs1/Xrs2) and is thought to result from different pathways. MRN is involved in a small scale resection process, while two redundant pathways depending on BLM and DNA2 on one hand, and on EXOI on another hand, would be involved in extensive resection (Mimitou and Symington 2008; Nimonkar, Genschel et al. 2011). In addition, processing ends involving a damaged nucleotide (resulting from chemical cleavage or from a bulk adduct), requires the CtIP/Sae2 protein together with RMN (Sartori, Lukas et al. 2007; Buis, Wu et al. 2008; Hartsuiker, Mizuno et al. 2009). Over-expression of the Trex2 exonuclease was shown to strongly stimulate imperfect NHEJ associated with loss of only a few base pairs (Bennardo, Gunn et al. 2009), while it inhibited various kinds of DNA repair events between distant sequences (such as Single-strand annealing, NHEJ between ends from different breaks, or NHEJ repair of a single DSB involving remote micro-homologies). In the same study, it was suggested that Trex2 did resect the 3' overhangs let by I-SceI in a non processive way. Thus, the type of stimulated pathway could in turn depend on the type of resection (length of resection, single strand vs. double strand, resection of 5' strand vs. 3' strand).

Thus, the efficiency of a compact TALEN, e.g. it ability to produce a desired event such as targeted mutagenesis or homologous gene targeting (see definition for full definition of "efficiency of compact TALEN"), can be enhanced by an enhancement or modification of its global DNA processing efficiency (see definition for full definition of "global DNA processing efficiency"), e.g. the global resultant or the overall result of different separate enzymatic activities that said compact TALEN.

In a preferred embodiment, enhancement of global DNA processing efficiency of a compact TALEN entity according to the present invention, compared to a starting compact TALEN entity, can be the hydrolysis of additional phosphodiester bonds at the cleavage site.

Said hydrolysis of additional phosphodiester bonds at the cleavage site by said at least one enhancer according to the invention can lead to different types of DSB resection affecting at said DSB cleavage site, one single DNA strand or both DNA strands, affecting either 5' overhangs ends, either 3' overhangs ends, or both ends and depending on the length of said resection. Thus, adding new

nickase or cleavase activities to the existing cleavase activity of a compact TALEN entity can enhance the efficiency of the resulting enhanced compact TALEN according to the invention, at a genomic locus of interest (Figure 8B-8E). As a non-limiting example, addition of two nickase activities on opposite strands (Figure 8D) or of a new cleavase activity generating a second DSB (Figure 8E) can result in a double-strand gap. As a consequence, perfect religation is no longer possible, and one or several alternative repair outcomes such as imprecise NHEJ, Homologous Recombination or SSA for instance, can be stimulated. As another non-limiting example, the addition of a single nickase activity can result in a single strand gap, and suppress the cohesiveness of the ends, which can also enhance the efficiency of the resulting enhanced compact TALEN at a genomic locus of interest, according to the invention, via stimulation of one or several alternative repair outcomes mentioned above.

In this aspect of the present invention, enhancement of DNA processing efficiency of a compact TALEN refers to the increase in the detected level of said DNA processing efficiency, against a target DNA sequence, of a enhanced compact TALEN in comparison to the activity of a first compact TALEN against the same target DNA sequence. Said first compact TALEN can be a starting compact TALEN, or a compact TALEN that has already been engineered or an enhanced compact TALEN according to the present invention. Several rounds of enhancement can be envisioned from a starting compact TALEN or from a starting enhanced compact TALEN.

In this aspect of the method of the present invention, enhancement of the DNA processing efficiency of the compact TALEN entity (or enhanced compact TALEN) refers to the increase in the detected level of said DNA processing efficiency against a target DNA sequence of interest or nearby said DNA sequence of interest in comparison to the efficiency of a first compact TALEN or starting compact TALEN against or nearby the same target DNA sequence. In this case, the starting compact TALEN is taken as the reference scaffold to measure the DNA processing efficiency. Said enhanced compact TALEN is an engineered compact TALEN comprising an enhancer domain according to this aspect of the invention. Said enhanced compact TALEN can also be taken as a reference scaffold for further enhancement of said DNA processing efficiency. As a non-limiting example, said DNA processing efficiency can result from a cleavage-induced recombination generated by said enhanced compact TALEN. In this case, said level of cleavage-induced recombination can be determined, for instance, by a cell-based recombination assay as described in the International PCT Application WO 2004/067736. Importantly, enhancement of efficacy in cells (enhanced generation of targeted mutagenesis or targeted recombination) can be, but is not necessarily associated with an

enhancement of the cleavage activity that could be detected in certain *in vitro* assays. For example, additional phosphodiesterase activities as described in Figure 8 could barely affect the cleavage profile, as detected by *in vitro* cleavage and separation of the cleavage products on an electrophoresis gel. However, as explained above, and in the legend of Figure 8, the DSB ends generated in this way could be more prone to induce detectable genomic rearrangements such as targeted mutagenesis (by imprecise NHEJ) or homologous recombination. Said enhancement in cleavage-induced recombination of said enhanced compact TALEN is at least a 5% enhancement compared to the starting scaffold or starting compact TALEN, more preferably at least a 10% enhancement, again more preferably at least a 15% enhancement, again more preferably at least a 20% enhancement, again more preferably at least a 25% enhancement, again more preferably a 50% enhancement, again more preferably an enhancement greater than 50%, resulting in an enhancement of DNA processing efficiency of said enhanced compact TALEN of at least 5% compared to the starting scaffold or starting compact TALEN, more preferably at least a 10% enhancement, again more preferably at least a 15% enhancement, again more preferably at least a 20% enhancement, again more preferably at least a 25% enhancement, again more preferably a 50% enhancement, again more preferably a enhancement greater than 50% .

In another preferred embodiment according to the method of the present invention, the peptidic linker that can link said enhancer domain to one part of said compact TALEN entity according to the method of the present invention can be selected from the group consisting of NFS1, NFS2, CFS1, RM2, BQY, QGPSG, LGPDGRKA, 1a8h_1, 1dnpA_1, 1d8cA_2, 1ckqA_3, 1sbp_1, 1ev7A_1, 1alo_3, 1amf_1, 1adjA_3, 1fcdC_1, 1al3_2, 1g3p_1, 1acc_3, 1ahjB_1, 1acc_1, 1af7_1, 1heiA_1, 1bia_2, 1igtB_1, 1nfkA_1, 1au7A_1, 1bpoB_1, 1b0pA_2, 1c05A_2, 1gcb_1, 1bt3A_1, 1b3oB_2, 16vpA_6, 1dhx_1, 1b8aA_1and 1qu6A_1 as listed in Table 3 (SEQ ID NO: 67 to SEQ ID NO: 104 and SEQ ID NO: 372 to SEQ ID NO: 415). In a more preferred embodiment, the peptidic linker that can said enhancer domain to one part of said compact TALEN entity according to the method of the present invention can be selected from the group consisting of NFS1 (SEQ ID NO: 98), NFS2 (SEQ ID NO: 99) and CFS1 (SEQ ID NO: 100). In the scope of the present invention is also encompassed the case where a peptidic linker is not needed to fuse one enhancer domain to one part of said compact TALEN entity in order to obtain a enhanced compact TALEN according to the present invention.

Depending from its structural composition [type of core TALE scaffold, type of catalytic domain(s) with associated enzymatic activities, type of enhancers and eventually type of linker(s)], a compact TALEN or an enhanced compact TALEN according to the present invention can comprise different levels of separate enzymatic activities able to differently process DNA as mentioned above. By adding new enzymatic activities to said compact TALEN or said enhanced compact TALEN or enhancing the DNA processing efficiency of one or several of its constitutive enzymatic activities, one can enhance the global DNA processing efficiency of one compact TALEN or enhanced compact TALEN in comparison to a starting compact TALEN or enhanced compact TALEN.

According to the present invention, compact TALENs are designed to alleviate the need for multiple independent protein moieties when targeting a DNA processing event. Importantly, the requisite "spacer" region and dual target sites essential for the function of current TALENs are unnecessary. As each end of the core TALE scaffold is amenable to fusion, the order (N- v.s C-terminal) of addition of the catalytic and enhancement domains can vary with the application. In addition, since the catalytic domain does not require specific DNA contacts, there are no restrictions on regions surrounding the core TALE scaffold, as non-limiting examples depicted in Figure 5: (A) N-terminal fusion construct to promote Homologous recombination induced by a cleavase domain or by a nickase domain. (B) C-terminal fusion construct with properties as in (A). (C) The attachment of two catalytic domains to both ends of the core TALE scaffold allows for dual cleavage with enhancement in NHEJ. Fusion junctions (N- vs. C-terminal) and linker designs can vary with the application.

According to the present invention, compact TALENs can be enhanced through the addition of a domain to promote existing or alternate activities as non-limiting examples depicted in figure 6 : (A) A standard compact TALEN with an enhancer domain fused to the C-terminus of its core TALE scaffold part. (B) The enhancer domain is fused to the compact TALEN via the N-terminus of its catalytic domain part. Such a configuration can be used to assist and/or anchor the catalytic domain part near the DNA to increase DNA processing activity. (C) The enhancer domain is sandwiched between the catalytic domain part and the core TALE scaffold part. The enhancer domain can promote communication between the flanking domains (i.e. to assist in catalysis and/or DNA binding) or can be used to overcome the requisite T nucleotide at position -1 of all TALE-based targets. (D) The enhancer domain is used to functionally replace the engineered core TALE scaffold N-terminal region. (E) The enhancer domain is used to functionally replace the engineered core TALE scaffold C-terminal region. Fusion junctions (N- vs. C-terminal) and linker designs can vary with the application.

According to the present invention, the nature of the catalytic domain(s) comprised in the compact TALEN and the enhanced compact TALEN is application dependent. As a non-limiting example, a nickase domain should allow for a higher HR/NHEJ ratio than a cleavase domain, thereby being more agreeable for therapeutic applications (McConnell Smith, Takeuchi et al. 2009; Metzger, McConnell-Smith et al. 2011). For example, the coupling of a cleavase domain on one side with a nickase domain on the other could result in excision of a single-strand of DNA spanning the binding region of a compact TALEN. The targeted generation of extended single-strand overhangs could be applied in applications that target DNA repair mechanisms. For targeted gene inactivation, the use of two cleavase domains is then preferred. In another preferred embodiment, the use of two nickase domains can be favored. Furthermore, the invention relates to a method for generating several distinct types of compact TALENs that can be applied to applications ranging from targeted DNA cleavage to targeted gene regulation.

In another aspect, the present invention relates to a compact TALEN comprising:

(i) One core TALE scaffold comprising different sets of Repeat Variable Dipeptide regions (RVDs) to change DNA binding specificity and target a specific single double-stranded DNA target sequence of interest, onto which a selection of catalytic domains can be attached to effect DNA processing;

(ii) At least one catalytic domain wherein said catalytic domain is capable of processing DNA nearby said single double-stranded DNA target sequence of interest when fused to said engineered core TALE scaffold from (i);

(iii) Optionally one peptidic linker to fuse said catalytic domain from (ii) to said engineered core TALE scaffold from (i) when needed;

such that said compact TALEN does not require dimerization to target a specific single double-stranded DNA target sequence of interest and process DNA nearby said single double-stranded DNA target sequence of interest. In other words, the compact TALEN according to the present invention is an active entity unit able, by itself, to target only one specific single double-stranded DNA target sequence of interest through one DNA binding domain and to process DNA nearby said single double-stranded DNA target sequence of interest.

The present invention relates to a compact TALEN monomer comprising:

(i) One core TALE scaffold comprising Repeat Variable Dipeptide regions (RVDs) having DNA binding specificity onto a specific double-stranded DNA target sequence of interest;

(ii) At least one catalytic domain wherein said catalytic domain is capable of processing DNA a few base pairs away from said double-stranded DNA target

sequence of interest when fused to the C or N terminal of said core TALE scaffold from (i);

(iii) Optionally one peptidic linker to fuse said catalytic domain from (ii) to said engineered core TALE scaffold from (i) when needed;

wherein said compact TALEN monomer is assembled to bind said target DNA sequence and process double-stranded DNA without requiring dimerization.

In another embodiment, said engineered core TALE scaffold of the compact TALEN according to the present invention comprises an additional N-terminal domain resulting in an engineered core TALE scaffold sequentially comprising a N-terminal domain and different sets of Repeat Variable Dipeptide regions (RVDs) to change DNA binding specificity and target a specific single double-stranded DNA target sequence of interest, onto which a selection of catalytic domains can be attached to effect DNA processing.

In another embodiment, said engineered core TALE scaffold of the compact TALEN according to the present invention comprises an additional C-terminal domain resulting in an engineered core TALE scaffold sequentially comprising different sets of Repeat Variable Dipeptide regions (RVDs) to change DNA binding specificity and target a specific single double-stranded DNA target sequence of interest and a C-terminal domain, onto which a selection of catalytic domains can be attached to effect DNA processing.

In another embodiment, said engineered core TALE-scaffold of the compact TALEN according to the present invention comprises additional N-terminus and a C-terminal domains resulting in an engineered core TALE scaffold sequentially comprising a N-terminal domain, different sets of Repeat Variable Dipeptide regions (RVDs) to change DNA binding specificity and target a specific single double-stranded DNA target sequence of interest and a C-terminal domain, onto which a selection of catalytic domains can be attached to effect DNA processing.

In another embodiment, said engineered core TALE-scaffold according to the present invention comprises the protein sequences selected from the group consisting of ST1 (SEQ ID NO: 134) and ST2 (SEQ ID NO: 135). In another embodiment, said engineered core TALE scaffold comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group consisting of SEQ ID NO: 134 and SEQ ID NO: 135. In another embodiment, said engineered core TALE-scaffold according to the present invention comprises the protein sequences selected from the group consisting of bT1-Avr (SEQ ID NO: 136), bT2-Avr (SEQ ID NO: 137), bT1-Pth (SEQ ID NO: 138) and bT2-Pth (SEQ ID NO: 139). In another embodiment, said engineered TALE-scaffold comprises a protein sequence having at least

80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group consisting of SEQ ID NO: 136 to SEQ ID NO: 139.

In a preferred embodiment, said additional N-terminus and C-terminal domains of engineered core TALE scaffold are derived from natural TALE. In a more preferred embodiment said additional N-terminus and C-terminal domains of engineered core TALE scaffold are derived from natural TALE selected from the group consisting of AvrBs3, PthXo1, AvrHah1, PthA, Tal1c as non-limiting examples. In another more preferred embodiment, said additional N-terminus and/or said C-terminal domains are truncated forms of respective N-terminus and/or said C-terminal domains of natural TALE like AvrBs3, PthXo1, AvrHah1, PthA, Tal1c as non-limiting examples, from which they are derived. In a more preferred embodiment, said additional N-terminus and C-terminal domains sequences of engineered core TALE scaffold are selected from the group consisting of ST1 SEQ ID NO: 134 and ST2 SEQ ID NO: 135 as respectively exemplified in baseline protein scaffolds bT1-Avr (SEQ ID NO: 136) or bT1-Pth (SEQ ID NO: 138) and bT2-Avr (SEQ ID NO: 137) or bT2-Pth (SEQ ID NO: 139).

In another embodiment, each RVD of said core scaffold is made of 30 to 42 amino acids, more preferably 33 or 34 wherein two critical amino acids located at positions 12 and 13 mediates the recognition of one nucleotide of said nucleic acid target sequence; equivalent two critical amino acids can be located at positions other than 12 and 13 specialy in RVDs taller than 33 or 34 amino acids long. Preferably, RVDs associated with recognition of the different nucleotides are HD for recognizing C, NG for recognizing T, NI for recognizing A, NN for recognizing G or A, NS for recognizing A, C, G or T, HG for recognizing T, IG for recognizing T, NK for recognizing G, HA for recognizing C, ND for recognizing C, HI for recognizing C, HN for recognizing G, NA for recognizing G, SN for recognizing G or A and YG for recognizing T, TL for recognizing A, VT for recognizing A or G and SW for recognizing A. More preferably, RVDs associated with recognition of the nucleotides C, T, A, G/A and G respectively are selected from the group consisting of NN or NK for recognizing G, HD for recognizing C, NG for recognizing T and NI for recognizing A, TL for recognizing A, VT for recognizing A or G and SW for recognizing A. In another embodiment, RVDS associated with recognition of the nucleotide C are selected from the group consisting of N* and RVDS associated with recognition of the nucleotide T are selected from the group consisting of N* and H*, where * denotes a gap in the repeat sequence that corresponds to a lack of amino acid residue at the second position of the RVD. In another embodiment, critical amino acids 12 and 13 can be mutated towards other amino acid residues in order to modulate their specificity towards nucleotides A, T, C and G and in particular to

enhance this specificity. By other amino acid residues is intended any of the twenty natural amino acid residues or unnatural amino acids derivatives.

In another embodiment, said core scaffold of the present invention comprises between 8 and 30 RVDs. More preferably, said core scaffold of the present invention comprises between 8 and 20 RVDs; again more preferably 15 RVDs.

In another embodiment, said core scaffold comprises an additional single truncated RVD made of 20 amino acids located at the C-terminus of said set of RVDs, i.e. an additional C-terminal half-RVD. In this case, said core scaffold of the present invention comprises between 8.5 and 30.5 RVDs, ".5" referring to previously mentioned half-RVD (or terminal RVD, or half-repeat). More preferably, said core scaffold of the present invention comprises between 8.5 and 20.5 RVDs, again more preferably, 15.5 RVDs. In a preferred embodiment, said half-RVD is in a core scaffold context which allows a lack of specificity of said half-RVD toward nucleotides A, C, G, T. In a more preferred embodiment, said half-RVD is absent.

In another embodiment, said core scaffold of the present invention comprises RVDs of different origins. In a preferred embodiment, said core scaffold comprises RVDs originating from different naturally occurring TAL effectors. In another preferred embodiment, internal structure of some RVDs of the core scaffold of the present invention are constituted by structures or sequences originated from different naturally occurring TAL effectors. In another embodiment, said core scaffold of the present invention comprises RVDs-like domains. RVDs-like domains have a sequence different from naturally occurring RVDs but have the same function and / or global structure within said core scaffold of the present invention.

In another embodiment, said additional N-terminal domain of said engineered core TALE scaffold of said compact TALEN according to the present invention is an enhancer domain. In another embodiment, said enhancer domain is selected from the group consisting of Puf RNA binding protein or Ankyrin super-family, as non-limiting examples. In another embodiment, said enhancer domain sequence is selected from the group consisting of protein domains of SEQ ID NO: 4 and SEQ ID NO: 5 as non-limiting examples listed in Table 1, a functional mutant, a variant or a derivative thereof. In another embodiment, said additional C-terminal domain of said engineered core TALE scaffold is an enhancer domain. In another embodiment, said enhancer domain is selected from the group consisting of hydrolase/transferase of Pseudomonas Aeuriginosa family, the polymerase domain

from the Mycobacterium tuberculosis Ligase D family, the initiation factor eIF2 from Pyrococcus family, the translation initiation factor Aif2 family as non-limiting examples. In another embodiment, said enhancer domain sequence is selected from the group consisting of protein domains of SEQ ID NO: 6 to SEQ ID NO: 9 as non-limiting examples listed in Table 1.

In another preferred embodiment, the catalytic domain that is capable of processing DNA nearby the single double-stranded DNA target sequence of interest, when fused to said engineered core TALE scaffold according to the present invention, is fused to the N-terminus part of said core TALE scaffold. In another preferred embodiment, said catalytic domain is fused to the C-terminus part of said core TALE scaffold. In another preferred embodiment two catalytic domains are fused to both N-terminus part of said core TALE scaffold and C-terminus part of said core TALE scaffold. In a more preferred embodiment, said catalytic domain has an enzymatic activity selected from the group consisting of nuclease activity, polymerase activity, kinase activity, phosphatase activity, methylase activity, topoisomerase activity, integrase activity, transposase activity or ligase activity. In another preferred embodiment, the catalytic domain fused to the core TALE scaffold of the present invention can be a transcription activator or repressor (i.e. a transcription regulator), or a protein that interacts with or modifies other proteins such as histones. Non-limiting examples of DNA processing activities of said compact TALEN of the present invention include, for example, creating or modifying epigenetic regulatory elements, making site-specific insertions, deletions, or repairs in DNA, controlling gene expression, and modifying chromatin structure.

In another more preferred embodiment, said catalytic domain has an endonuclease activity. In another more preferred embodiment, said catalytic domain of the compact TALEN according to the present invention has cleavage activity on said double-stranded DNA according to the method of the present invention. In another more preferred embodiment, said catalytic domain has a nickase activity on said double-stranded DNA according to the method of the present invention. In another more preferred embodiment, said catalytic domain is selected from the group consisting of proteins MmeI, Colicin-E7 (CEA7_ECOLX), Colicin-E9, APFL, EndA, Endo I (END1_ECOLI), Human Endo G (NUCG_HUMAN), Bovine Endo G (NUCG_BOVIN), R.HinP1I, I-BasI, I-BmoI, I-HmuI, I-TevI, I-TevII, I-TevIII, I-TwoI, R.MspI, R.MvaI, NucA, NucM, Vvn, Vvn_CLS, Staphylococcal nuclease (NUC_STAAU), Staphylococcal nuclease (NUC_STAHY), Micrococcal nuclease (NUC_SHIFL), Endonuclease yncB, Endodeoxyribonuclease I (ENRN_BPT7), Metnase, Nb.BsrDI, BsrDI A, Nt.BspD6I (R.BspD6I large subunit), ss.BspD6I (R.BspD6I small subunit), R.PleI, MlyI, AlwI, Mva1269I, BsrI, BsmI, Nb.BtsCI, Nt.BtsCI, R1.BtsI, R2.BtsI, BbvCI subunit 1, BbvCI subunit 2, Bpu10I alpha subunit, Bpu10I beta subunit, BmrI, BfiI, I-CreI, hExoI (EXO1_HUMAN), Yeast ExoI (EXO1_YEAST), E.coli ExoI, Human

TREX2, Mouse TREX1, Human TREX1, Bovine TREX1, Rat TREX1, Human DNA2, Yeast DNA2 (DNA2_YEAST) and VP16, as listed in Table 2 (SEQ ID NO: 10 to SEQ ID NO: 66 and SEQ ID NO: 1, 366 & 367), a functional mutant, a variant or a derivative thereof. In another preferred embodiment said catalytic domain of the compact TALEN according to the present invention is I-TevI (SEQ ID NO: 20), a functional mutant, a variant or a derivative thereof. In another preferred embodiment, catalytic domain I-TevI (SEQ ID NO: 20), a functional mutant, a variant or a derivative thereof is fused to the N-terminal domain of said core TALE scaffold according to the compact TALEN of the present invention. In another preferred embodiment, said compact TALEN according to the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group of SEQ ID NO: 426-432.

In another preferred embodiment, said catalytic domain of the compact TALEN according to the present invention is ColE7 (SEQ ID NO: 11), a functional mutant, a variant or a derivative thereof. In another preferred embodiment, catalytic domain ColE7 (SEQ ID NO: 11), a functional mutant, a variant or a derivative thereof is fused to the N-terminal domain of said core TALE scaffold according to the method of the present invention. In another preferred embodiment, catalytic domain ColE7 (SEQ ID NO: 11), a functional mutant, a variant or a derivative thereof is fused to the C-terminal domain of said core TALE scaffold according to the method of the present invention. In another preferred embodiment, said compact TALEN according to the method of the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group of SEQ ID NO: 435-438.

In another preferred embodiment, said catalytic domain of the compact TALEN according to the present invention is NucA (SEQ ID NO: 26), a functional mutant, a variant or a derivative thereof. In another preferred embodiment, catalytic domain NucA (SEQ ID NO: 26), a functional mutant, a variant or a derivative thereof is fused to the N-terminal domain of said core TALE scaffold according to the method of the present invention. In another preferred embodiment, catalytic domain NucA (SEQ ID NO: 26), a functional mutant, a variant or a derivative thereof is fused to the C-terminal domain of said core TALE scaffold according to the method of the present invention. In another preferred embodiment, said compact TALEN according to the method of the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group of SEQ ID NO: 433-434.

In another preferred embodiment, said catalytic domain is I-CreI (SEQ ID NO: 1), a functional mutant, a variant or a derivative thereof. In another preferred embodiment, catalytic domain I-CreI (SEQ ID NO: 1), a functional mutant, a variant or a derivative thereof is fused to the N-terminal domain of said core TALE scaffold according to the method of the present invention. In another preferred embodiment, catalytic domain I-CreI (SEQ ID NO: 1), a functional mutant, a variant or a derivative thereof is fused to the C-terminal domain of said core TALE scaffold according to the present invention. In another preferred embodiment, said compact TALEN according to the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group of SEQ ID NO: 439-441 and SEQ ID NO: 444-446.

In another embodiment, said catalytic domain is a restriction enzyme such as MmeI, R-HinPII, R.MspI, R.MvaI, Nb.BsrDI, BsrDI A, Nt.BspD6I, ss.BspD6I, R.PleI, MlyI and AlwI as non-limiting examples listed in table 2. In another more preferred embodiment, said catalytic domain has an exonuclease activity. In another more preferred embodiment, any combinations of two catalytic domains selected from the group consisting of proteins MmeI, Colicin-E7 (CEA7_ECOLX), Colicin-E9, APFL, EndA, Endo I (END1_ECOLI), Human Endo G (NUCG_HUMAN), Bovine Endo G (NUCG_BOVIN), R.HinP1I, I-BasI, I-BmoI, I-HmuI, I-TevI, I-TevII, I-TevIII, I-TwoI, R.MspI, R.MvaI, NucA, NucM, Vvn, Vvn_CLS, Staphylococcal nuclease (NUC_STAAU), Staphylococcal nuclease (NUC_STAHY), Micrococcal nuclease (NUC_SHIFL), Endonuclease yncB, Endodeoxyribonuclease I (ENRN_BPT7), Metnase, Nb.BsrDI, BsrDI A, Nt.BspD6I (R.BspD6I large subunit), ss.BspD6I (R.BspD6I small subunit), R.PleI, MlyI, AlwI, Mva1269I, BsrI, BsmI, Nb.BtsCI, Nt.BtsCI, R1.BtsI, R2.BtsI, BbvCI subunit 1, BbvCI subunit 2, Bpu10I alpha subunit, Bpu10I beta subunit, BmrI, BfiI, I-CreI, hExoI (EXO1_HUMAN), Yeast ExoI (EXO1_YEAST), E.coli ExoI, Human TREX2, Mouse TREX1, Human TREX1, Bovine TREX1, Rat TREX1, Human DNA2, Yeast DNA2 (DNA2_YEAST) and VP16, as listed in Table 2 (SEQ ID NO: 10 to SEQ ID NO: 66 and SEQ ID NO: 1, 366 & 367), a functional mutant, a variant or a derivative of these protein domains thereof, can be fused to both N-terminus part and C-terminus part of said core TALE scaffold, respectively. For example, I-HmuI catalytic domain can be fused to the N-terminus part of said core TALE scaffold and ColE7 catalytic domain can be fused to the C-terminus part of said core TALE scaffold. In another example, I-TevI catalytic domain can be fused to the N-terminus part of said core TALE scaffold and ColE7 catalytic domain can be fused to the C-terminus part of said core TALE scaffold.

Table 14 below gives non-limiting examples of combinations of catalytic domains that can be comprised in the compact TALEN monomer according to the present invention. Optionally, FokI (SEQ

ID NO :368) can be used in combination with another catalytic domain according to the list of Table2.

Catalytic domain fused to N-terminus part of core TALE scaffold	Catalytic domain fused to C-terminus part of core TALE scaffold	Dual-cleavage TALENS
I-TevI	I-TevI	TevI-TevI
ColE7	ColE7	ColE7-ColE7
NucA	NucA	NucA-NucA
I-TevI	ColE7	TevI-ColE7
I-TevI	NucA	TevI-NucA
ColE7	I-TevI	ColE7-TevI
ColE7	NucA	ColE7-NucA
NucA	I-TevI	NucA-TevI
NucA	ColE7	NucA-ColE7

Table 14: Examples of combinations of catalytic domains respectively fused to N and C-terminus part of compact TALEN core scaffolds according to the present invention leading to dual-cleavage TALENs.

In a preferred embodiment according to the present invention, said unique compact TALEN monomer comprises a combination of two catalytic domains respectively fused to the N-terminus part and to the C-terminus part of said core TALE scaffold selected from the group consisting of:

(i) A Nuc A domain (SEQ ID NO: 26) in N-terminus and a Nuc A domain (SEQ ID NO: 26) in C-terminus;

(ii) A ColE7 domain (SEQ ID NO: 11) in N-terminus and a ColE7 domain (SEQ ID NO: 11) in C-terminus;

(iii) A TevI domain (SEQ ID NO: 20) in N-terminus and a ColE7 domain (SEQ ID NO: 11) in C-terminus;

(iv) A TevI domain (SEQ ID NO: 20) in N-terminus and a NucA domain (SEQ ID NO: 26) in C-terminus;

(v) A ColE7 domain (SEQ ID NO: 11) in N-terminus and a NucA domain (SEQ ID NO: 26) in C-terminus;

(vi) A NucA domain (SEQ ID NO: 26) in N-terminus and a ColE7 domain (SEQ ID NO: 11) in C-terminus.

In another preferred embodiment, said compact TALEN according to the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group consisting of SEQ ID NO: 448 and 450.

In another preferred embodiment, said compact TALEN according to the present invention comprises a combination of two catalytic domains respectively fused to the C-terminus part and to the N-terminus part of said core TALE scaffold selected from the group consisting of:

(i) A TevI domain (SEQ ID NO: 20) in N-terminus and a FokI domain (SEQ ID NO: 368) in C-terminus;

(ii) A TevI domain (SEQ ID NO: 20) in N-terminus and a TevI domain (SEQ ID NO: 20) in C-terminus;

(iii) A scTrex2 domain (SEQ ID NO: 451) in N-terminus and a FokI domain (SEQ ID NO: 368) in C-terminus.

In another preferred embodiment, said compact TALEN according to the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group consisting of SEQ ID NO: 447-450 and SEQ ID NO: 452.

In the scope of the present invention, it can be envisioned to insert said catalytic domain and / or said enhancer domain between two parts of the engineered core TALE scaffold according to the invention, each part comprising one set of RVDs. In this last case, the number of RVDs for each part of the engineered core TALE scaffold can be the same or not. In other words, it can be envisioned to split said core TALE scaffold of the present invention to insert one catalytic domain and / or one enhancer domain between the resulting two parts of said engineered core TALE scaffold. In another preferred embodiment, said compact TALEN according to the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group consisting of SEQ ID NO: 453-455.

In other words, the compact TALEN monomer of the present invention comprises a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group consisting of SEQ ID NO: 420-450 and 452-455.

In another preferred embodiment according to the method of the present invention, the peptidic linker that can link said catalytic domain to the core TALE scaffold according to the method of the present invention can be selected from the group consisting of NFS1, NFS2, CFS1, RM2, BQY, QGPSG, LGPDGRKA, 1a8h_1, 1dnpA_1, 1d8cA_2, 1ckqA_3, 1sbp_1, 1ev7A_1, 1alo_3, 1amf_1, 1adjA_3, 1fcdC_1, 1al3_2, 1g3p_1, 1acc_3, 1ahjB_1, 1acc_1, 1af7_1, 1heiA_1, 1bia_2, 1igtB_1, 1nfkA_1, 1au7A_1, 1bpoB_1, 1b0pA_2, 1c05A_2, 1gcb_1, 1bt3A_1, 1b3oB_2, 16vpA_6, 1dhx_1, 1b8aA_1and 1qu6A_1, as listed in Table 3 (SEQ ID NO: 67 to SEQ ID NO: 104 and SEQ ID NO: 372 to SEQ ID NO: 415). In a more preferred embodiment, the peptidic linker that can link said catalytic domain to the core TALE scaffold according to the method of the present invention can be selected from the group consisting of NFS1 (SEQ ID NO: 98), NFS2 (SEQ ID NO: 99) andCFS1 (SEQ ID NO: 100). In the scope of the present invention is also encompassed the case where a peptidic linker is not needed to fuse a catalytical domain to the TALE scaffold in order to obtain a cTALEN according to the present invention.

Depending from its structural composition [type of core TALE scaffold, type of catalytic domain(s) with associated enzymatic activities and eventually type of linker(s)], a compact TALEN according to the present invention can comprise different levels of separate enzymatic activities able to differently process DNA, resulting in a global DNA processing efficiency for said compact TALEN, each one of said different enzymatic activities having their own DNA processing efficiency.

In another preferred embodiment, the compact TALEN according to the present invention further comprises:

(i) at least one enhancer domain;

(ii) Optionally one peptide linker to fuse said enhancer domain to one part of said compact TALEN active entity;

thereby obtaining a compact TALEN entity with enhanced DNA processing efficiency nearby a single double-stranded DNA target sequence of interest, i.e an enhanced compact TALEN.

In other words, said unique compact TALEN monomer further comprises:

(i) At least one enhancer domain;

(ii) Optionally one peptide linker to fuse said enhancer domain to one part of said unique compact TALEN monomer active entity.

In another more preferred embodiment, said enhancer domain is fused to N-terminus of the core TALE scaffold part of said compact TALEN entity. In another more preferred embodiment, said enhancer domain is fused to C-terminus of the core TALE scaffold part of said compact TALEN entity. In another more preferred embodiment, said enhancer domain is fused to the catalytic domain part of said compact TALEN entity. In another more preferred embodiment, said enhancer domain is fused between the N-terminus part of the core TALE scaffold and the catalytic part of said compact TALEN entity. In another more preferred embodiment, said enhancer domain is fused between the C-terminus part of the core TALE scaffold and the catalytic part of said compact TALEN entity. In the scope of the present invention, it can be envisioned to insert said catalytic domain and / or enhancer domain between two parts of the engineered core TALE scaffold according to the invention, each part comprising one set of RVDs. In this last case, the number of RVDs for each engineered core TALE scaffold can be the same or not. In other words, it can be envisioned to split said core TALE scaffold of the present invention to insert one catalytic domain and / or one enhancer domain between the resulting two parts of said engineered core TALE scaffold.

In another preferred embodiment, said enhancer domain is catalytically active or not, providing functional and/or structural support to said compact TALEN entity. In a more preferred embodiment, said enhancer domain consists of a protein domain selected from the group consisting of MmeI, Colicin-E7 (CEA7_ECOLX), Colicin-E9, APFL, EndA, Endo I (END1_ECOLI), Human Endo G (NUCG_HUMAN), Bovine Endo G (NUCG_BOVIN), R.HinP1I, I-BasI, I-BmoI, I-HmuI, I-TevI, I-TevII, I-TevIII, I-TwoI, R.MspI, R.MvaI, NucA, NucM, Vvn, Vvn_CLS, Staphylococcal nuclease (NUC_STAAU), Staphylococcal nuclease (NUC_STAHY), Micrococcal nuclease (NUC_SHIFL), Endonuclease yncB, Endodeoxyribonuclease I (ENRN_BPT7), Metnase, Nb.BsrDI, BsrDI A, Nt.BspD6I (R.BspD6I large subunit), ss.BspD6I (R.BspD6I small subunit), R.PleI, MlyI, AlwI, Mva1269I, BsrI, BsmI, Nb.BtsCI, Nt.BtsCI, R1.BtsI, R2.BtsI, BbvCI subunit 1, BbvCI subunit 2, Bpu10I alpha subunit, Bpu10I beta subunit, BmrI, BfiI, I-CreI, hExoI (EXO1_HUMAN), Yeast ExoI (EXO1_YEAST), E.coli ExoI, Human TREX2, Mouse TREX1, Human TREX1, Bovine TREX1, Rat TREX1, Human DNA2, Yeast DNA2 (DNA2_YEAST) and VP16, as listed in Table 2 (SEQ ID NO: 10 to SEQ ID NO: 66 and SEQ ID NO: 1, 366 & 367), a functional mutant, a variant or a derivative of these protein domains thereof. In another more preferred embodiment, said enhancer domain consists of a catalytically active derivative of the protein domains listed above and in Table 2, providing functional and/or structural support to said

compact TALEN entity. In another preferred embodiment, said enhancer domain consists of a catalytically inactive derivative of the protein domains listed above and in Table 2, providing structural support to said compact TALEN entity. In another preferred embodiment, said enhancer domain is selected from the group consisting of I-TevI (SEQ ID NO: 20), ColE7 (SEQ ID NO: 11) and NucA (SEQ ID NO: 26).

In a more preferred embodiment, said enhanced compact TALEN according to the present invention can comprise a second enhancer domain. In this embodiment, said second enhancer domain can have the same characteristics than the first enhancer domain. In a more preferred embodiment, said second enhancer domain provides structural support to enhanced compact TALEN entity. In another more preferred embodiment, said second enhancer domain provides functional support to enhanced compact TALEN entity. In a more preferred embodiment, said second enhancer domain provides structural and functional supports to enhanced compact TALEN entity. In a more preferred embodiment, said enhanced compact TALEN entity comprises one catalytic domain and one enhancer domain. In another more preferred embodiment said enhanced compact TALEN entity comprises one catalytic domain and two enhancer domains. In another more preferred embodiment said enhanced compact TALEN entity comprises two catalytic domains and one enhancer domains. In another more preferred embodiment said enhanced compact TALEN entity comprises two catalytic domains and two enhancer domains.

In a more preferred embodiment, said second enhancer domain consists of a protein domain derived from a protein selected from the group consisting of MmeI, Colicin-E7 (CEA7_ECOLX), Colicin-E9, APFL, EndA, Endo I (END1_ECOLI), Human Endo G (NUCG_HUMAN), Bovine Endo G (NUCG_BOVIN), R.HinP1I, I-BasI, I-BmoI, I-HmuI, I-TevI, I-TevII, I-TevIII, I-TwoI, R.MspI, R.MvaI, NucA, NucM, Vvn, Vvn_CLS, Staphylococcal nuclease (NUC_STAAU), Staphylococcal nuclease (NUC_STAHY), Micrococcal nuclease (NUC_SHIFL), Endonuclease yncB, Endodeoxyribonuclease I (ENRN_BPT7), Metnase, Nb.BsrDI, BsrDI A, Nt.BspD6I (R.BspD6I large subunit), ss.BspD6I (R.BspD6I small subunit), R.PleI, MlyI, AlwI, Mva1269I, BsrI, BsmI, Nb.BtsCI, Nt.BtsCI, R1.BtsI, R2.BtsI, BbvCI subunit 1, BbvCI subunit 2, Bpu10I alpha subunit, Bpu10I beta subunit, BmrI, BfiI, I-CreI, hExoI (EXO1_HUMAN), Yeast ExoI (EXO1_YEAST), E.coli ExoI, Human TREX2, Mouse TREX1, Human TREX1, Bovine TREX1, Rat TREX1, Human DNA2, Yeast DNA2 (DNA2_YEAST) and VP16, as listed in Table 2 (SEQ ID NO: 10 to SEQ ID NO: 66 and SEQ ID NO: 1, 366 & 367), a functional mutant, a variant or a derivative of these protein domains thereof . In another more preferred embodiment, said second enhancer domain consists of a catalytically active derivative of the protein domains listed above and in Table 2, providing functional and/or structural support to said enhanced compact TALEN entity. In another preferred embodiment, said second enhancer domain consists of a catalytically inactive derivative of

the protein domains listed above and in Table 2, providing structural support to said enhanced compact TALEN entity.

In another more preferred embodiment, any combinations of catalytic and/or enhancer domains listed above, as non-limiting examples, can be envisioned to be fused to said core TALE scaffold providing structural and/or functional support to said compact TALEN entity. More preferably, combinations of catalytic domains listed in Table 14. Again more preferably, combinations of catalytic domains selected from the group of TevI (SEQ ID NO: 20), ColE7 (SEQ ID NO: 11) and NucA (SEQ ID NO: 26) can be envisioned. Optionally, FokI (SEQ ID NO: 368) can be used in combination with another catalytic domain according to the list of Table2. Such combinations of catalytic and/or enhancer domains can be envisioned regarding the envisioned applications for using the method of the present invention.

Depending from its structural composition [type of core TALE scaffold, type of catalytic domain(s) with associated enzymatic activities, type of linker(s) and type of enhancer(s) domains], an enhanced compact TALEN according to the present invention can present different levels of separate enzymatic activities able to differently process DNA, resulting in a global DNA processing efficiency for said enhanced compact TALEN, each one of said different enzymatic activities having their own DNA processing efficiency.

In this preferred embodiment, the DNA processing efficiency of the compact TALEN entity according to the present invention can be enhanced by the engineering of at least one enhancer domain and one peptidic linker thereby obtaining a compact TALEN entity with enhanced DNA processing activity nearby a single double-stranded DNA target sequence of interest, i.e a enhanced compact TALEN according to the present invention.

Depending from its structural composition, the global DNA processing efficiency that is enhanced in said enhanced compact TALEN according to the present invention, can have a dominant enzymatic activity selected from the group consisting of a nuclease activity, a polymerase activity, a kinase activity, a phosphatase activity, a methylase activity, a topoisomerase activity, an integrase activity, a transposase activity or a ligase activity as non-limiting examples. In a more preferred embodiment, the global DNA processing efficiency that is enhanced in said enhanced compact TALEN according to the present invention is a combination of different enzymatic activities selected from the group consisting of a nuclease activity, a polymerase activity, a kinase activity, a phosphatase activity, a methylase activity, a topoisomerase activity, an integrase activity, a transposase activity or a ligase

activity as non-limiting examples. In a more preferred embodiment, the global DNA processing efficiency that is enhanced in said enhanced compact TALEN according to the present invention is one of its different enzymatic activities selected from the group consisting of a nuclease activity, a polymerase activity, a kinase activity, a phosphatase activity, a methylase activity, a topoisomerase activity, an integrase activity, a transposase activity or a ligase activity as non-limiting examples. In this case, the global DNA processing efficiency is equivalent to one DNA processing activity amongst the enzymatic activities mentioned above. In another more preferred embodiment, said DNA processing activity of the compact TALEN entity which is enhanced by the enhancer is a cleavase activity or a nickase activity or a combination of both a cleavase activity and a nickase activity.

Enhancement of DNA processing efficiency of a compact TALEN entity according to the present invention can be a consequence of a structural support by said at least one enhancer domain. In a preferred embodiment, said structural support enhances the binding of a compact TALEN entity according to the invention for said DNA target sequence compared to the binding of a starting compact TALEN entity for the same DNA target sequence, thereby indirectly assisting the catalytic domain(s) to obtain a compact TALEN entity with enhanced DNA processing activity. In another preferred embodiment, said structural support enhances the existing catalytical activity of a compact TALEN entity for a DNA target sequence compared to the binding of a starting compact TALEN entity for the same DNA target sequence to obtain a compact TALEN entity with enhanced DNA processing activity.
In another preferred embodiment, said enhancer according to the present invention both enhances the binding of the compact TALEN entity for said DNA target sequence and the catalytic activity of the catalytic domain(s) to obtain a compact TALEN entity with enhanced DNA processing activity. All these non-limiting examples lead to a compact TALEN entity with enhanced DNA processing efficiency for a DNA target sequence at a genomic locus of interest, i.e an enhanced compact TALEN according to the present invention.

Enhancement of DNA processing efficiency of a compact TALEN entity according to the present invention, compared to a starting compact TALEN entity, can also be a consequence of a fuctional support by said at least one enhancer domain. In a preferred embodiment, said functional support can be the consequence of the hydrolysis of additional phosphodiester bonds. In a more preferred embodiment, said functional support can be the hydrolysis of additional phosphodiester bonds by a protein domain derived from a nuclease. In a more preferred embodiment, said functional support can be the hydrolysis of additional phosphodiester bonds by a protein domain derived from an

endonuclease. In a more preferred embodiment, said functional support can be the hydrolysis of additional phosphodiester bonds by a protein domain derived from a cleavase. In another more preferred embodiment, said functional support can be the hydrolysis of additional phosphodiester bonds by a protein domain derived from a nickase. In a more preferred embodiment, said functional support can be the hydrolysis of additional phosphodiester bonds by a protein domain derived from an exonuclease.

In genome engineering experiments, the efficiency of rare-cutting endonuclease, e.g. their ability to induce a desired event (Homologous gene targeting, targeted mutagenesis, sequence removal or excision) at a locus, depends on several parameters, including the specific activity of the nuclease, probably the accessibility of the target, and the efficacy and outcome of the repair pathway(s) resulting in the desired event (homologous repair for gene targeting, NHEJ pathways for targeted mutagenesis).

Cleavage by peptidic rare cutting endonucleases usually generates cohesive ends, with 3' overhangs for LAGLIDADG meganucleases (Chevalier and Stoddard 2001) and 5' overhangs for Zinc Finger Nucleases (Smith, Bibikova et al. 2000). These ends, which result from hydrolysis of phosphodiester bonds, can be re-ligated in vivo by NHEJ in a seamless way (i.e a scarless re-ligation). The restoration of a cleavable target sequence allows for a new cleavage event by the same endonuclease, and thus, a series of futile cycles of cleavage and re-ligation events can take place. Indirect evidences have shown that even in the yeast Saccharomyces cerevisiae, such cycles could take place upon continuous cleavage by the HO endonuclease (Lee, Paques et al. 1999). In mammalian cells, several experiment have shown that perfect re-ligation of compatible cohesive ends resulting from two independent but close I-SceI-induced DSBs is an efficient process (Guirouilh-Barbat, Huck et al. 2004; Guirouilh-Barbat, Rass et al. 2007; Bennardo, Cheng et al. 2008; Bennardo, Gunn et al. 2009). Absence of the Ku DNA repair protein does not significantly affect the overall frequency of NHEJ events rejoining the ends from the two DSBs; however it very strongly enhances the contribution of imprecise NHEJ to the repair process in CHO immortalized cells and mouse ES cells (Guirouilh-Barbat, Huck et al. 2004; Guirouilh-Barbat, Rass et al. 2007; Bennardo, Cheng et al. 2008). Furthermore, the absence of Ku stimulates I-SceI-induced events such as imprecise NHEJ (Bennardo, Cheng et al. 2008), single-strand annealing (Bennardo, Cheng et al. 2008) and gene conversion (Pierce, Hu et al. 2001; Bennardo, Cheng et al. 2008) in mouse ES cells . Similar observations shave been made with cells deficient for the XRCC4 repair protein (Pierce, Hu et al. 2001; Guirouilh-Barbat, Rass et al. 2007; Bennardo, Gunn et al. 2009) (although XRCC4 deficiency affects the overal level of NHEJ in CHO cells (Guirouilh-Barbat, Rass et al. 2007)) or for DNA-PK (Pierce, Hu et al. 2001). In contrast, knock-down

of CtIP has been shown to suppresses "alt-NHEJ" (a Ku- and XRCC4-independent form of NHEJ more prone to result in imprecise NHEJ), single-strand annealing and gene conversion, while not affecting the overall level of rejoining of two compatible ends generated by I-SceI (Bennardo, Cheng et al. 2008). Thus, competition between different DSB repair pathways can affect the spectrum or repair events resulting from a nuclease-induced DSB.

In addition, DSB resection is important for certain DSB pathways. Extensive DSB resection, resulting in the generation of large single stranded regions (a few hundred nucleotides at least), has been shown in yeast to initiate single strand annealing (Sugawara and Haber 1992) and strand invasion, the ATP-dependant step that initiates many homologous recombination events of DNA duplex invasion by an homologous strand that (White and Haber 1990; Sun, Treco et al. 1991) (for a review of mechanisms, see (Paques and Haber 1999)). In eukaryotic cells DSB resection depends on several proteins including BLM/Sgs1 and DNA2, EXOI, and the MRN complex (Mre11, Rad50, Nbs1/Xrs2) and is thought to result from different pathways. MRN is involved in a small scale resection process, while two redundant pathways depending on BLM and DNA2 on one hand, and on EXOI on another hand, would be involved in extensive resection (Mimitou and Symington 2008; Nimonkar, Genschel et al. 2011). In addition, processing ends involving a damaged nucleotide (resulting from chemical cleavage or from a bulk adduct), requires the CtIP/Sae2 protein together with RMN (Sartori, Lukas et al. 2007; Buis, Wu et al. 2008; Hartsuiker, Mizuno et al. 2009). Over-expression of the Trex2 exonuclease was shown to strongly stimulate imperfect NHEJ associated with loss of only a few base pairs (Bennardo, Gunn et al. 2009), while it inhibited various kinds of DNA repair events between distant sequences (such as Single-strand annealing, NHEJ between ends from different breaks, or NHEJ repair of a single DSB involving remote micro-homologies). In the same study, it was suggested that Trex2 did resect the 3' overhangs let by I-SceI in a non processive way. Thus, the type of stimulated pathway could in turn depend on the type of resection (length of resection, single strand vs. double strand, resection of 5' strand vs. 3' strand).

Thus, the efficiency of a compact TALEN, e.g. it ability to produce a desired event such as targeted mutagenesis or homologous gene targeting (see definition for full definition of "efficiency of compact TALEN"), can be enhanced by an enhancement or modification of its global DNA processing efficiency (see definition for full definition of "global DNA processing efficiency"), e.g. the global resultant or the overall result of different separate enzymatic activities that said compact TALEN.

In a preferred embodiment, enhancement of global DNA processing efficiency of a compact TALEN entity according to the present invention, compared to a starting compact TALEN entity, can be the hydrolysis of additional phosphodiester bonds at the cleavage site.

Said hydrolysis of additional phosphodiester bonds at the cleavage site by said at least one enhancer according to the invention can lead to different types of DSB resection affecting at said DSB cleavage site, one single DNA strand or both DNA strands, affecting either 5' overhangs ends, either 3' overhangs ends, or both ends and depending on the length of said resection. Thus, adding new nickase or cleavase activities to the existing cleavase activity of a compact TALEN entity can enhance the efficiency of the resulting enhanced compact TALEN according to the invention, at a genomic locus of interest (Figure 8B-8E). As a non-limiting example, addition of two nickase activities on opposite strands (Figure 8D) or of a new cleavase activity generating a second DSB (Figure 8E) can result in a double-strand gap. As a consequence, perfect religation is not possible anymore, and one or several alternative repair outcomes such as imprecise NHEJ, Homologous Recombination or SSA for instance, can be stimulated. As another non-limiting example, the addition of a single nickase activity can result in a single strand gap, and suppress the cohesivity of the ends, which can also enhances the efficiency of the resulting enhanced compact TALEN at a genomic locus of interest, according to the invention, via stimulation of one or several alternative repair outcomes mentioned above.

In this aspect of the present invention, enhancement of DNA processing efficiency of a compact TALEN refers to the increase in the detected level of said DNA processing efficiency, against a target DNA sequence, of a compact TALEN in comparison to the activity of a first compact TALEN against the same target DNA sequence. Said first compact TALEN can be a starting compact TALEN, or a compact TALEN that has already been engineered or an enhanced compact TALEN according to the present invention. Several rounds of enhancement can be envisioned from a starting compact TALEN or from a starting enhanced compact TALEN.

In this aspect of the present invention, enhancement of the DNA processing efficiency of the compact TALEN entity (or enhanced compact TALEN) refers to the increase in the detected level of said DNA processing efficiency against a target DNA sequence of interest or nearby said DNA sequence of interest in comparison to the efficiency of a first compact TALEN or starting compact TALEN against or nearby the same target DNA sequence. In this case, the starting compact TALEN is taken as the reference scaffold to measure the DNA processing efficiency. Said enhanced compact TALEN is an engineered compact TALEN comprising an enhancer domain according to this aspect of

the invention. Said enhanced compact TALEN can also be taken as a reference scaffold for further enhancement in said DNA processing efficiency. As a non-limiting example, said DNA processing efficiency can result from a cleavage-induced recombination generated by said enhanced compact TALEN. In this case, said level of cleavage-induced recombination can be determined, for instance, by a cell-based recombination assay as described in the International PCT Application WO 2004/067736. Importantly, enhancement of efficacy in cells (enhanced generation of targeted mutagenesis or targeted recombination) can be, but is not necessarily associated with an enhancement of the cleavage activity that could be detected in certain *in vitro* assays. For example, additional phosphodiesterase activities as described in Figure 8 could barely affect the cleavage profile, as detected by *in vitro* cleavage and separation of the cleavage products on an electrophoresis gel. However, as explained above, and in the legend of Figure 8, the DSB ends generated in this way could be more prone to induce detectable genomic rearrangements such as targeted mutagenesis (by imprecise NHEJ) or homologous recombination. Said enhancement in cleavage-induced recombination of said enhanced compact TALEN is at least a 5% enhancement compared to the starting scaffold or starting compact TALEN, more preferably at least a 10% enhancement, again more preferably at least a 15% enhancement, again more preferably at least a 20% enhancement, again more preferably at least a 25% enhancement, again more preferably a 50% enhancement, again more preferably a enhancement greater than 50%, resulting in an enhancement of DNA processing efficiency of said enhanced compact TALEN of at least 5% compared to the starting scaffold or starting compact TALEN, more preferably at least a 10% enhancement, again more preferably at least a 15% enhancement, again more preferably at least a 20% enhancement, again more preferably at least a 25% enhancement, again more preferably a 50% enhancement, again more preferably a enhancement greater than 50% .

In another preferred embodiment according to the method of the present invention, the peptidic linker that can link said enhancer domain to one part of said compact TALEN entity according to the method of the present invention can be selected from the group consisting of NFS1, NFS2, CFS1, RM2, BQY, QGPSG, LGPDGRKA, 1a8h_1, 1dnpA_1, 1d8cA_2, 1ckqA_3, 1sbp_1, 1ev7A_1, 1alo_3, 1amf_1, 1adjA_3, 1fcdC_1, 1al3_2, 1g3p_1, 1acc_3, 1ahjB_1, 1acc_1, 1af7_1, 1heiA_1, 1bia_2, 1igtB_1, 1nfkA_1, 1au7A_1, 1bpoB_1, 1b0pA_2, 1c05A_2, 1gcb_1, 1bt3A_1, 1b3oB_2, 16vpA_6, 1dhx_1, 1b8aA_1and 1qu6A_1 as listed in table 3 (SEQ ID NO: 67 to SEQ ID NO: 104 and SEQ ID NO: 372 to SEQ ID NO: 415). In a more preferred embodiment, the peptidic linker that can said enhancer domain to one part of said compact TALEN entity according to the method of the present invention can be selected from the group consisting of NFS1 (SEQ ID NO: 98), NFS2 (SEQ ID NO: 99) and CFS1

(SEQ ID NO: 100). In the scope of the present invention is also encompassed the case where a peptidic linker is not needed to fuse one enhancer domain to one part of said compact TALEN entity in order to obtain a enhanced compact TALEN according to the present invention.

Depending from its structural composition [type of core TALE scaffold, type of catalytic domain(s) with associated enzymatic activities, type of enhancers and eventually type of linker(s)], a compact TALEN or an enhanced compact TALEN according to the present invention can comprise different levels of separate enzymatic activities able to differently process DNA as mentioned above. By adding new enzymatic activities to said compact TALEN or enhanced compact TALEN or enhancing the DNA processing efficiency of one or several of its constitutive enzymatic activities, one can enhance the global DNA processing efficiency of one compact TALEN or enhanced compact TALEN in comparison to a starting compact TALEN or enhanced compact TALEN.

According to the present invention, compact TALENs are designed to alleviate the need for multiple independent protein moieties when targeting a DNA processing event. Importantly, the requisite "spacer" region and dual target sites essential for the function of current TALENs are unnecessary, as compact TALENs according to the invention comprises a core TALE scaffold containing only one DNA binding domain to target a specific single double-stranded DNA target sequence of interest and process DNA nearby said single double-stranded DNA target sequence of interest. As each end of the core TALE scaffold is amenable to fusion, the order (N- v.s C-terminal) of addition of the catalytic and enhancement domains can vary with the application. In addition, since the catalytic domain does not require specific DNA contacts, there are no restrictions on regions surrounding the core TALE scaffold, as non-limiting examples depicted in Figure 5: (A) N-terminal fusion construct to promote Homologous recombination induced by a cleavase domain or by a nickase domain. (B) C-terminal fusion construct with properties as in (A). (C) The attachment of two catalytic domains to both ends of the core TALE scaffold allows for dual cleavage with enhancement in NHEJ. Fusion junctions (N- vs. C-terminal) and linker designs can vary with the application.

According to the present invention, compact TALENs can be enhanced through the addition of a domain to promote existing or alternate activities as non-limiting examples depicted in figure 6 : (A) A standard compact TALEN with an enhancer domain fused to the C-terminus of its core TALE scaffold part. (B) The enhancer domain is fused to the compact TALEN via the N-terminus of its catalytic domain part. Such a configuration can be used to assist and/or anchor the catalytic domain

part near the DNA to increase DNA processing activity. (C) The enhancer domain is sandwiched between the catalytic domain part and the core TALE scaffold part. The enhancer domain can promote communication between the flanking domains (i.e. to assist in catalysis and/or DNA binding) or can be used to overcome the requisite T nucleotide at position -1 of all TALE-based targets. (D) The enhancer domain is used to functionally replace the engineered core TALE scaffold N-terminal region. (E) The enhancer domain is used to functionally replace the engineered core TALE scaffold C-terminal region. Fusion junctions (N- vs. C-terminal) and linker designs can vary with the application.

According to the present invention, the nature of the catalytic domain(s) comprised in the compact TALEN and the enhanced compact TALEN is application dependent. As a non-limiting example, a nickase domain should allow for a higher HR/NHEJ ratio than a cleavase domain, thereby being more agreeable for therapeutic applications (McConnell Smith, Takeuchi et al. 2009; Metzger, McConnell-Smith et al. 2011). . For example, the coupling of a cleavase domain on one side with a nickase domain on the other could result in excision of a single-strand of DNA spanning the binding region of a compact TALEN. The targeted generation of extended single-strand overhangs could be applied in applications that target DNA repair mechanisms. For targeted gene inactivation, the use of two cleavase domains is then preferred. In another preferred embodiment, the use of two nickase domains can be favored. Furthermore, the invention relates to a method for generating several distinct types of compact TALENs that can be applied to applications ranging from targeted DNA cleavage to targeted gene regulation.

The present invention also relates to methods for use of said compact TALENs according to the invention for various applications ranging from targeted DNA cleavage to targeted gene regulation. In a preferred embodiment, the present invention relates to a method for increasing targeted HR (and NHEJ) when Double-Strand break activity is promoted in a compact TALEN targeting a DNA target sequence according to the invention. In another more preferred embodiment, the addition of at least two catalytically active cleavase enhancer domains according to the invention allows to increase Double-strand break-induced mutagenesis by leading to a loss of genetic information and preventing any scarless re-ligation of targeted genomic locus of interest by NHEJ.

In another preferred embodiment, the present invention relates to a method for increasing targeted HR with less NHEJ (i.e in a more conservative fashion) when Single-Strand Break activity is promoted in a compact TALEN targeting a DNA target sequence according to the invention.

In another preferred embodiment, the present invention relates to a method for increasing excision of a single-strand of DNA spanning the binding region of a compact TALEN when both one cleavase enhancer domain and one nickase enhancer domain, respectively, are fused to both N-terminus and C-terminus of a core TALE scaffold according to the invention.

In another preferred embodiment, the present invention relates to a method for treatment of a genetic disease caused by a mutation in a specific single double-stranded DNA target sequence in a gene, comprising administering to a subject in need thereof an effective amount of a variant of a compact TALEN according to the present invention.

In another preferred embodiment, the present invention relates to a method for inserting a transgene into a specific single double-stranded DNA target sequence of a genomic locus of a cell, tissue or non-human animal, or a plant wherein at least one compact TALEN of the present invention is transitory or not introduced into said cell, tissue, non-human animal or plant.

In another embodiment, the present invention relates to a method to modulate the activity of a compact TALEN when expressed in a cell wherein said method comprises the step of introducing in said cell an auxiliary domain modulating the activity of said compact TALEN. In a preferred embodiment, the present invention relates to a method which allows to have a temporal control of activity of a compact TALEN when expressed in a cell by introducing in said cell an auxiliary domain modulating the activity of said compact TALEN once said compact TALEN achieved its activity (DNA cleavage, DNA nicking or other DNA processing activities). In a preferred embodiment, the present invention relates to a method to inhibit the activity of a compact TALEN when expressed in a cell wherein said method comprises the step of introducing in said cell an auxiliary domain inhibiting the activity of said compact TALEN. In a more preferred embodiment, the catalytic domain of said compact TALEN is NucA (SEQ ID NO: 26) and said auxiliary domain is NuiA (SEQ ID NO: 229), a functional mutant, a variant or a derivative thereof. In another more preferred embodiment, the catalytic domain of said compact TALEN is ColE7 (SEQ ID NO: 11) and said auxiliary domain is Im7 (SEQ ID NO: 230), a functional mutant, a variant or a derivative thereof.

Is also encompassed in the scope of the present invention a recombinant polynucleotide encoding a compact TALEN, a dual compact TALEN, or an enhanced compact TALEN according to the present invention.Is also encompassed in the scope of the present invention, a vector comprising a recombinant polynucleotide encoding for a compact TALEN or an enhanced compact TALEN according to the present invention.

Is also encompassed in the scope of the present invention, a host cell which comprises a vector and /or a recombinant polynucleotide encoding for a compact TALEN or an enhanced compact TALEN according to the present invention.

Is also encompassed in the scope of the present invention, a non-human transgenic animal comprising a vector and /or a recombinant polynucleotide encoding for a compact TALEN or an enhanced compact TALEN according to the present invention.

Is also encompassed in the scope of the present invention, a transgenic plant comprising a vector and /or a recombinant polynucleotide encoding for a compact TALEN or an enhanced compact TALEN according to the present invention.

The present invention also relates to kits used to implement the method according to the present invention. More preferably, is encompassed in the scope of the present invention, a kit comprising a compact TALEN or an enhanced compact TALEN according to the present invention and instructions for use said kit in enhancing DNA processing efficiency of a single double-stranded DNA target sequence of interest.

For purposes of therapy, the compact TALENs of the present invention and a pharmaceutically acceptable excipient are administered in a therapeutically effective amount. Such a combination is said to be administered in a "therapeutically effective amount" if the amount administered is physiologically significant. An agent is physiologically significant if its presence results in a detectable change in the physiology of the recipient. In the present context, an agent is physiologically significant if its presence results in a decrease in the severity of one or more symptoms of the targeted disease and in a genome correction of the lesion or abnormality. Vectors comprising targeting DNA and/or nucleic acid encoding a compact TALEN can be introduced into a cell by a variety of methods (e.g., injection, direct uptake, projectile bombardment, liposomes, electroporation). Compact TALENs can be stably or transiently expressed into cells using expression vectors. Techniques of expression in eukaryotic cells are well known to those in the art. (See Current Protocols in Human Genetics: Chapter 12 "Vectors For Gene Therapy" & Chapter 13 "Delivery Systems for Gene Therapy").

In one further aspect of the present invention, the compact TALEN of the present invention is substantially non-immunogenic, i.e., engender little or no adverse immunological response. A variety of methods for ameliorating or eliminating deleterious immunological reactions of this sort can be used in accordance with the invention. In a preferred embodiment, the compact TALEN is substantially free of N-formyl methionine. Another way to avoid unwanted immunological reactions is to conjugate compact TALEN to polyethylene glycol ("PEG") or polypropylene glycol ("PPG") (preferably of 500 to 20,000 daltons average molecular weight (MW)). Conjugation with PEG or PPG, as described by Davis *et al.* (US 4,179,337) for example, can provide non-immunogenic, physiologically active, water soluble compact TALEN conjugates with anti-viral activity. Similar methods also using a polyethylene--polypropylene glycol copolymer are described in Saifer *et al.* (US 5,006,333).

In another aspect of the present invention is a composition comprising a compact TALEN or an enhanced compact TALEN according to the present invention and a carrier. More preferably , is a pharmaceutical composition comprising a compact TALEN or an enhanced compact TALEN according to the present invention and a pharmaceutically active carrier known in the state of the art.

In the scope of the present invention and for all the applications mentioned above, it can be envisioned to use more than one compact TALEN (i.e one compact TALEN active entity) or more than one enhanced compact TALENs (i.e one enhanced compact TALEN active entity) for DNA processing according to the invention. In a preferred embodiment, two different compact TALENs or two enhanced compact TALENs can be used. In this embodiment, as non-limiting examples, said two different compact TALENs can comprise the same core TALE scaffold or not; said two different compact TALENs can comprise the same set of Repeat Variable Dipeptides or not; said two different compact TALENs can comprise the same catalytic domain or not. When two identical compact TALENs active entities are used for DNA processing according to the invention, they can be considered as a homodimeric pair of compact TALENs active entities. When two non identical compact TALENs active entities are used for DNA processing according to the invention, they can be considered as a heterodimeric pair of compact TALENs active entities. As non-limiting example, when two compact TALEN according to the present invention are used, one of the compact TALEN can modulate the activity of the other one, leading for instance to an enhanced DNA processing event compared to the same DNA processing event achieved by only one compact TALEN; in this non-limiting example, a Trans-TALEN modulates and enhances the catalytic activity of an initial compact TALEN.

In another preferred embodiment, three compact TALENs or three enhanced compact TALENs can be used. In another preferred embodiment, more than three compact TALENs or three enhanced compact TALENs can be used for DNA processing according to the invention. In another preferred embodiment, a combination of compact TALENs and enhanced compact TALENs can be used for DNA processing according to the invention. As a non-limiting example, one compact TALEN and one enhanced compact TALEN can be used. As another non-limiting example, one compact TALEN and one dual-cleavage compact TALEN can be used. In another preferred embodiment, a combination of compact TALENs, enhanced compact TALENs and dual-cleavage compact TALENs can be used, said compact TALENs comprising the same catalytic domain or not, the same core TALE scaffold or not. When several compact TALENs have to be used, DNA target sequence for each compact TALENs of the combination to be used can be located on a same endogenous genomic DNA locus of interest or not. Said DNA target sequences can be located at an approximative distance of 1000 base pairs (bps). More preferably, said DNA target sequences can be located at an approximative distance of 500 bps or 200 bps, or 100 bps, or 50 bps, or 20 bps, 19 bps, 18 bps, 17 bps, 16 bps, 15 bps, 14 bps, 13 bps, 12 bps, 11 bps, 10 bps, 9 bps, 8 bps, 7 bps, 6 bps, 5 bps, 4 bps, 3 bps, 2 bps, 1 bp. Said DNA target sequences located at distances mentioned above are "nearby" DNA sequences in reference to the target DNA sequence for DNA processing according to the present invention.

In another preferred embodiment, two compact TALENs active entities can be used as a way of achieving two different DNA processing activities nearby a DNA target sequence according to the invention. As a non-limiting example, two compact TALENs targeting said DNA sequence or DNA sequences nearby said targeted DNA sequence and comprising each one a nickase-derived catalytic domain can be used; in this case, this use of two compact TALENs active entities can represent an alternative way of achieving a Double Strand Break nearby a said DNA target sequence, compared to the use of one compact TALEN targeting said DNA sequence and comprising a cleavase-derived catalytic domain, or not. As another non-limiting example, one compact TALEN comprising a cleavase-derived domain and one compact TALEN comprising an exonuclease-derived domain can be used to make a Double Strand Break and create a gap, respectively, to achieve an imprecise NHEJ event at the genomic locus of interest comprising said DNA target sequence. In this case, even if each compact TALEN forming this heterodimeric pair of compact TALENs is active by itself, each of these active entities is dependent of the other one to achieve the wanted resulting DNA processing activity. Indeed, in this particular case, the wanted resulting activity is a gap created by the exonuclease activity, said exonuclease activity being possible only from the Double Strand Break achieved by the cleavase domain of the other compact TALENs. In the scope of the present

invention, is also envisioned the case where two identical compact TALEN active entities (a homodimeric pair of compact TALENs) are dependent each other to achieve a wanted resulting DNA processing activity.

Definitions

- Amino acid residues in a polypeptide sequence are designated herein according to the one-letter code, in which, for example, Q means Gln or Glutamine residue, R means Arg or Arginine residue and D means Asp or Aspartic acid residue.

- Amino acid substitution means the replacement of one amino acid residue with another, for instance the replacement of an Arginine residue with a Glutamine residue in a peptide sequence is an amino acid substitution.

- Enhanced/increased/improved DNA processing activity, refers to an increase in the detected level of a given compact TALEN or enhanced compact TALEN associated DNA processing activity against a target DNA sequence or DNA target sequence by a second compact TALEN or enhanced compact TALEN in comparison to the activity of a first compact TALEN or enhanced compact TALEN against the target DNA sequence. The second compact TALEN or enhanced compact TALEN can be a variant of the first one and can comprise one or more substituted amino acid residues in comparison to the first compact TALEN or enhanced compact TALEN. The second compact TALEN or enhanced compact TALEN can be a variant of the first one and can comprise one or more catalytic and/or enhancer domains in comparison to said first compact TALEN or enhanced compact TALEN. This definition more broadly applies for other endonucleases and rare-cutting endonucleases.

- DNA processing activity refers to a particular / given enzymatic activity of said compact TALEN or enhanced compact TALEN or more broadly to qualify the enzymatic activity of a rare-cutting endonuclease. Said DNA processing activity can refer to a cleavage activity, either a cleavase activity either a nickase activity, more broadly a nuclease activity but also a polymerase activity, a kinase activity, a phosphatase activity, a methylase activity, a topoisomerase activity, an integrase activity, a transposase activity or a ligase activity as non-limiting examples. In the scope of this definition, said given DNA processing activity of a particular enzymatic activity can also be described as DNA processing efficiency of said particular enzymatic activity. Methods for enhancing compact TALEN or enhanced compact TALEN DNA processing activity according to this definition are encompassed in the present invention.

- Global DNA processing efficiency describes, for a compact TALEN or an enhanced compact TALEN according to the present invention, the global resultant or the overall effect of different separate enzymatic activities that said compact TALEN or enhanced compact TALEN can comprise. According to these different separate enzymatic activities, a compact TALEN or an enhanced compact TALEN presents a global capacity to process DNA nearby a target sequence in a genomic locus of interest, i.e a global DNA processing efficiency. Said global DNA processing efficiency can qualify or rank a second given compact TALEN or enhanced compact TALEN in comparison to a first given compact TALEN or enhanced compact TALEN. Depending on said compact TALENs or enhanced compact TALENs structural composition [type of core TALE scaffold, type of catalytic domain(s) with associated enzymatic activities, eventually type of linker(s) and type of enhancer(s) domains], said global DNA processing efficiency can refer to only one enzymatic activity, two enzymatic activities, three enzymatic activities, four enzymatic activities or more than four enzymatic activities. Said global DNA processing efficiency can refer to the sum of individual enzymatic activities. Said global DNA processing efficiency can refer to the synergy or combined effect of different enzymatic activities comprised in a given compact TALEN or enhanced compact TALEN. An enhancement of the DNA processing efficiency of a compact TALEN according to the present invention can reflect a synergy, an enhanced combined effect, resulting in an enhanced compact TALEN characterized by a global DNA processing efficiency that is greater than the sum of respective DNA processing efficiencies of separate starting compact TALEN or than the sum of respective DNA processing efficiencies of separate enzymatic activities comprised in a same starting compact TALEN.

- Efficiency of a rare-cutting endonuclease according to the present invention is the property for said rare-cutting endonuclease of producing a desired event. This desired event can be for example Homologous gene targeting, targeted mutagenesis, or sequence removal or excision. The efficiency of the desired event depends on several parameters, including the specific activity of the nuclease and the repair pathway(s) resulting in the desired event (efficacy of homologous repair for gene targeting, efficacy and outcome of NHEJ pathways for targeted mutagenesis). Efficiency of a rare cutting endonuclease for a locus is intended to mean its ability to produce a desired event at this locus. Efficiency of a rare cutting endonuclease for a target is intended to mean its ability to produce a desired event as a consequence of cleavage of this target.

- Nucleotides are designated as follows: one-letter code is used for designating the base of a nucleoside: a is adenine, t is thymine, c is cytosine, and g is guanine. For the degenerated nucleotides, r represents g or a (purine nucleotides), k represents g or t, s represents g or c, w represents a or t, m represents a or c, y represents t or c (pyrimidine nucleotides), d represents g, a or t, v represents g, a or c, b represents g, t or c, h represents a, t or c, and n represents g, a, t or c.

- by "meganuclease", is intended a rare-cutting endonuclease subtype having a double-stranded DNA target sequence greater than 12 bp. Said meganuclease is either a dimeric enzyme, wherein each domain is on a monomer or a monomeric enzyme comprising the two domains on a single polypeptide.

- by "meganuclease domain" is intended the region which interacts with one half of the DNA target of a meganuclease and is able to associate with the other domain of the same meganuclease which interacts with the other half of the DNA target to form a functional meganuclease able to cleave said DNA target.

- by "endonuclease variant", "rare-cutting endonuclease variant", "chimeric rare-cutting endonuclease variant" or "meganuclease variant" or "compact TALEN variant", or "enhanced compact TALEN variant" or "dual cleavage compact TALEN variant" or "variant" it is intended an endonuclease, rare-cutting endonuclease, chimeric rare-cutting endonuclease, meganuclease, or compact TALEN, enhanced compact TALEN, dual cleavage compact TALEN obtained by replacement of at least one residue in the amino acid sequence of the parent endonuclease, rare-cutting endonuclease, chimeric rare-cutting endonuclease, meganuclease or compact TALEN, enhanced compact TALEN, dual cleavage compact TALEN with at least a different amino acid."Variant" designation also applies for instance for an enhanced compact TALEN comprising at least one supplementary protein domain (catalytic or enhancer domain) in comparison to the starting compact TALEN entity. Are also encompassed in the scope of the present definition, variants and protein domains comprised in these variants which present a sequence with high percentage of identity or high percentage of homology with sequences of compact TALENs, enhanced compact TALENs, dual-cleavage compact TALENs or protein domains and polypeptides according to the present invention, at nucleotidic or polypeptidic levels. By high percentage of identity or high percentage of homology it is intended 60%, more preferably 70%, more preferably 75%, more preferably 80%, more preferably 85%, more preferably 90%, more preferably 95, more preferably 97%, more preferably 99% or any integer comprised between 60% and 99%.

- by "peptide linker", "peptidic linker" or "peptide spacer" it is intended to mean a peptide sequence which allows the connection of different monomers in a fusion protein and the adoption of the correct conformation for said fusion protein activity and which does not alter the specificity of either of the monomers for their targets. Peptide linkers can be of various sizes, from 3 amino acids to 50 amino acids as a non limiting indicative range. Peptide linkers can also be structured or unstructured.

- by "related to", particularly in the expression "one cell type related to the chosen cell type or organism", is intended a cell type or an organism sharing characteristics with said chosen cell type or said chosen organism; this cell type or organism related to the chosen cell type or organism, can be derived from said chosen cell type or organism or not.

- by "subdomain" it is intended the region of a LAGLIDADG homing endonuclease core domain which interacts with a distinct part of a homing endonuclease DNA target half-site.

- by "targeting DNA construct/minimal repair matrix/repair matrix" it is intended to mean a DNA construct comprising a first and second portion that are homologous to regions 5' and 3' of the DNA target *in situ*. The DNA construct also comprises a third portion positioned between the first and second portion which comprise some homology with the corresponding DNA sequence *in situ* or alternatively comprise no homology with the regions 5' and 3' of the DNA target *in situ*. Following cleavage of the DNA target, a homologous recombination event is stimulated between the genome containing the targeted gene comprised in the locus of interest and the repair matrix, wherein the genomic sequence containing the DNA target is replaced by the third portion of the repair matrix and a variable part of the first and second portions of the repair matrix.

- by "functional variant" is intended a catalytically active variant of a protein, such variant can have additional properties compared to its parent protein. As a non-limiting example, a functional variant of a meganuclease can be able to cleave a DNA target sequence, preferably said target being a new target which is not cleaved by the parent meganuclease. This definition also applies to compact TALENs, enhanced compact TALENs, dual-cleavage compact TALENs or protein domains that constitute such TALENs according to the present invention. Are also encompassed in the scope of the present definition, functional variants, polypeptides and protein domains comprised in these molecules which present a sequence with high percentage of identity or high percentage of homology with sequences of compact TALENs, enhanced compact TALENs, dual-cleavage compact TALENs or protein domains and polypeptides according to the present invention, at nucleotidic or polypeptidic levels. By high percentage of identity or high percentage of homology it is intended 60%, more preferably 70%, more preferably 75%, more preferably 80%, more preferably 85%, more preferably 90%, more preferably 95, more preferably 97%, more preferably 99% or any integer comprised between 60% and 99%.

- by "derived from" or "derivative(s)" it is intended to mean for instance a meganuclease variant which is created from a parent meganuclease and hence the peptide sequence of the meganuclease

variant is related to (primary sequence level) but derived from (mutations) the peptide sequence of the parent meganuclease. In this definition, mutations encompass deletions or insertions of several amino acid residues; as non-limiting example, a truncated variant of an I-CreI meganuclease is considered as a scaffold derived from I-CreI meganuclease. This expression can also apply to compact TALENs, enhanced compact TALENs, dual-cleavage compact TALENs or protein domains that constitute such TALENs according to the present invention. Are also encompassed in the scope of the present definition, derivatives of compact TALENs, enhanced compact TALENs, dual-cleavage compact TALENs or protein domains and derivatives of polypeptides according to the present invention which present a sequence with high percentage of identity or high percentage of homology with sequences of compact TALENs, enhanced compact TALENs, dual-cleavage compact TALENs or protein domains and polypeptides according to the present invention, at nucleotidic or polypeptidic levels. By high percentage of identity or high percentage of homology it is intended 60%, more preferably 70%, more preferably 75%, more preferably 80%, more preferably 85%, more preferably 90%, more preferably 95, more preferably 97%, more preferably 99% or any integer comprised between 60% and 99%.

- by "I-*Cre*I" is intended the wild-type I-*Cre*I having the sequence of pdb accession code 1g9y, corresponding to the sequence SEQ ID NO: 1 in the sequence listing. In the present Patent Application, *I-Cre*I variants described can comprise an additional Alanine after the first Methionine of the wild type *I-Cre*I sequence (SEQ ID NO: 1). These variants may also comprise two additional Alanine residues and an Aspartic Acid residue after the final Proline of the wild type *I-Cre*I sequence as shown in SEQ ID NO: 106. These additional residues do not affect the properties of the enzyme and to avoid confusion these additional residues do not affect the numeration of the residues in *I-Cre*I or a variant referred in the present Patent Application, as these references exclusively refer to residues of the wild type *I-Cre*I enzyme (SEQ ID NO: 1) as present in the variant, so for instance residue 2 of *I-Cre*I is in fact residue 3 of a variant which comprises an additional Alanine after the first Methionine.

- by compact TALEN, enhanced compact TALEN, dual-cleavage compact TALEN with novel specificity is intended a variant of these proteins having a pattern of cleaved targets different from that of their respective parent compact TALENs, enhanced compact TALENs, dual-cleavage compact TALENs. The terms "novel specificity", "modified specificity", "novel cleavage specificity", "novel substrate specificity" which are equivalent and used indifferently, refer to the specificity of the variant towards the nucleotides of the DNA target sequence.

- by "I-*Cre*I site" is intended a 22 to 24 bp double-stranded DNA sequence which is cleaved by I-*Cre*I. I-*Cre*I sites include the wild-type non-palindromic I-*Cre*I homing site and the derived palindromic sequences such as the sequence 5'- $t_{-12}c_{-11}a_{-10}a_{-9}a_{-8}a_{-7}c_{-6}g_{-5}t_{-4}c_{-3}g_{-2}t_{-1}a_{+1}c_{+2}g_{+3}a_{+4}c_{+5}g_{+6}t_{+7}t_{+8}t_{+9}t_{+10}g_{+11}a_{+12}$ (SEQ ID NO: 2), also called C1221 or C1221 target.

- by "domain" or "core domain" is intended the "LAGLIDADG homing endonuclease core domain" which is the characteristic $\alpha\beta\beta\alpha\beta\beta\alpha$ fold of the homing endonucleases of the LAGLIDADG family, corresponding to a sequence of about one hundred amino acid residues. Said domain comprises four beta-strands ($\beta_1\beta_2\beta_3\beta_4$) folded in an anti-parallel beta-sheet which interacts with one half of the DNA target. This domain is able to associate with another LAGLIDADG homing endonuclease core domain which interacts with the other half of the DNA target to form a functional endonuclease able to cleave said DNA target. For example, in the case of the dimeric homing endonuclease I-*Cre*I (163 amino acids), the LAGLIDADG homing endonuclease core domain corresponds to the residues 6 to 94.

- by "beta-hairpin" is intended two consecutive beta-strands of the antiparallel beta-sheet of a LAGLIDADG homing endonuclease core domain ($\beta_1\beta_2$ or $\beta_3\beta_4$) which are connected by a loop or a turn.

- by "single-chain meganuclease", "single-chain chimeric meganuclease", "single-chain meganuclease derivative", "single-chain chimeric meganuclease derivative" or "single-chain derivative" is intended a meganuclease comprising two LAGLIDADG homing endonuclease domains or core domains linked by a peptidic spacer. The single-chain meganuclease is able to cleave a chimeric DNA target sequence comprising one different half of each parent meganuclease target sequence.

- by "DNA target", "DNA target sequence", "target DNA sequence", "target sequence", "target-site", "target", "site", "site of interest", "recognition site", "polynucleotide recognition site", "recognition sequence", "homing recognition site", "homing site", "cleavage site" is intended a double-stranded palindromic, partially palindromic (pseudo-palindromic) or non-palindromic polynucleotide sequence that is recognized and can be cleaved by a LAGLIDADG homing endonuclease such as I-*Cre*I, or a variant, or a single-chain chimeric meganuclease derived from I-*Cre*I. Said DNA target sequence can be qualified as "cleavable" by an endonuclease, rare-cutting endonuclease, chimeric rare-cutting endonuclease or meganuclease when recognized within a genomic sequence and known to correspond to the DNA target sequence of a given endonuclease, rare-cutting endonuclease, chimeric rare-cutting endonuclease or meganuclease or a variant of such

endonuclease, rare-cutting endonuclease, chimeric rare-cutting endonuclease or meganuclease. These terms refer to a specific DNA location, preferably a genomic location, but also a portion of genetic material that can exist independently to the main body of genetic material such as plasmids, episomes, virus, transposons or in organelles such as mitochondria or chloroplasts as non-limiting examples, at which a double stranded break (cleavage) can be induced by the endonuclease, rare-cutting endonuclease, chimeric rare-cutting endonuclease or meganuclease. For the LAGLIDADG subfamily of rare-cutting endonucleases, the DNA target is defined by the 5' to 3' sequence of one strand of the double-stranded polynucleotide, as indicate above for C1221 (SEQ ID NO: 2). Cleavage of the DNA target can occur at the nucleotides at positions +2 and -2, respectively for the sense and the antisense strand. Unless otherwise indicated, the position at which cleavage of the DNA target by an I-*Cre*I-derived variant can occur, corresponds to the cleavage site on the sense strand of the DNA target. In the particular case of compact TALENs, a subclass of chimeric rare-cutting endonucleases, the following expressions "DNA target", "DNA target sequence", "target DNA sequence", "target sequence" , "target-site", "target" , "site", "site of interest", "recognition site", "polynucleotide recognition site", and "recognition sequence" can apply to qualify their specific DNA target sequence with the particularity that said specific DNA target sequence recognized by the compact TALEN according to the invention is the one or not that is processed and/or cut by the compact TALEN. A compact TALEN, an enhanced compact TALEN or a dual cleavage compact TALEN according to the present invention can process and / or cut DNA within said specific DNA target sequence. A compact TALEN, an enhanced compact TALEN or a dual cleavage compact TALEN can also process and / or cut DNA outside said specific DNA target sequence.

- By "DNA nearby said specific DNA target sequence" or by "DNA nearby" is intended DNA sequence or sequences located within or outside said specific DNA target sequence. Are also intended DNA sequence or sequences bound by a compact TALEN or an enhanced compact TALEN at said specific DNA target sequence location or DNA located at a 5' or 3' distance of 1-100 base pairs (bps), 1-50 base pairs (bps) or 1-25 base pairs (bps) from said specific DNA target sequence.

When several compact TALENs have to be used in a particular genome engineering application, DNA target sequence for each compact TALENs of the combination to be used can be located on a same endogenous genomic DNA locus of interest or not. Said DNA target sequences can be located at an approximative distance of 1-1000 base pairs (bps), more preferably 1-500 bps, more preferably 1-100 bps, more preferably 1-100 bps, more preferably 1-50 bps, more preferably 1-25 bps, more preferably 1-10 bps. In another embodiment, said DNA target sequence for each compact TALENs of the combination to be used can be located on the same DNA strand or not. Said DNA target

sequences located at distances mentioned above are "nearby" DNA sequences in reference to the target DNA sequence for DNA processing according to the present invention.

-by "single double-stranded DNA target sequence" is intended a compact-TALEN or enhanced compact TALEN or dual-cleavage compact TALEN binding site. The recognition DNA binding site of a compact-TALEN or enhanced compact TALEN or dual-cleavage compact TALEN can be ranging from 12 to 100 base pairs (bp) in length, usually greater than 12 bps in length.

- by "DNA target half-site", "half cleavage site" or half-site" is intended the portion of the DNA target which is bound by each LAGLIDADG homing endonuclease core domain.

- The term "endonuclease" refers to any wild-type or variant enzyme capable of catalyzing the hydrolysis (cleavage) of bonds between nucleic acids within a DNA or RNA molecule, preferably a DNA molecule. Endonucleases can be classified as rare-cutting endonucleases when having typically a polynucleotide recognition greater than 12 base pairs (bp) in length, more preferably of 14-45 bp. Rare-cutting endonucleases significantly increase HR by inducing DNA double-strand breaks (DSBs) at a defined locus (Rouet, Smih et al. 1994; Rouet, Smih et al. 1994; Choulika, Perrin et al. 1995; Pingoud and Silva 2007). Rare-cutting endonucleases can for example be a homing endonuclease (Paques and Duchateau 2007), a chimeric Zinc-Finger nuclease (ZFN) resulting from the fusion of engineered zinc-finger domains with the catalytic domain of a restriction enzyme such as FokI (Porteus and Carroll 2005) or a chemical endonuclease (Eisenschmidt, Lanio et al. 2005 ; Arimondo, Thomas et al. 2006; Simon, Cannata et al. 2008). In chemical endonucleases, a chemical or peptidic cleaver is conjugated either to a polymer of nucleic acids or to another DNA recognizing a specific target sequence, thereby targeting the cleavage activity to a specific sequence. Chemical endonucleases also encompass synthetic nucleases like conjugates of orthophenanthroline, a DNA cleaving molecule, and triplex-forming oligonucleotides (TFOs), known to bind specific DNA sequences (Kalish and Glazer 2005). Such chemical endonucleases are comprised in the term "endonuclease" according to the present invention.
Rare-cutting endonucleases can also be for example TALENs, a new class of chimeric nucleases using a FokI catalytic domain and a DNA binding domain derived from Transcription Activator Like Effector (TALE), a family of proteins used in the infection process by plant pathogens of the Xanthomonas genus (Boch, Scholze et al. 2009; Boch, Scholze et al. 2009; Moscou and Bogdanove 2009; Moscou and Bogdanove 2009; Christian, Cermak et al. 2010; Christian, Cermak et al. 2010; Li, Huang et al. 2010; Li, Huang et al. 2011). The functional layout of a FokI-based TALE-nuclease (TALEN) is

essentially that of a ZFN, with the Zinc-finger DNA binding domain being replaced by the TALE domain. As such, DNA cleavage by a TALEN requires two DNA recognition regions flanking an unspecific central region. Rare-cutting endonucleases encompassed in the present invention can also be derived from TALENs. The authors of the present invention have developed a new type of TALENs that can be engineered to specifically recognize and process target DNA efficiently. These novel "compact TALENs" (cTALENs) do not require dimerization for DNA processing activity, thereby alleviating the need for "dual" target sites with intervening DNA "spacers"; these compact TALENs can be seen as one subclass of rare-cutting endonucleases or chimeric rare-cutting endonucleases according to the present invention.

Rare-cutting endonuclease can be a homing endonuclease, also known under the name of meganuclease. Such homing endonucleases are well-known to the art (Stoddard 2005). Homing endonucleases recognize a DNA target sequence and generate a single- or double-strand break. Homing endonucleases are highly specific, recognizing DNA target sites ranging from 12 to 45 base pairs (bp) in length, usually ranging from 14 to 40 bp in length. The homing endonuclease according to the invention may for example correspond to a LAGLIDADG endonuclease, to a HNH endonuclease, or to a GIY-YIG endonuclease.

In the wild, meganucleases are essentially represented by homing endonucleases. Homing Endonucleases (HEs) are a widespread family of natural meganucleases including hundreds of proteins families (Chevalier and Stoddard 2001). These proteins are encoded by mobile genetic elements which propagate by a process called "homing": the endonuclease cleaves a cognate allele from which the mobile element is absent, thereby stimulating a homologous recombination event that duplicates the mobile DNA into the recipient locus. Given their exceptional cleavage properties in terms of efficacy and specificity, they could represent ideal scaffolds to derive novel, highly specific endonucleases.

HEs belong to four major families. The LAGLIDADG family, named after a conserved peptidic motif involved in the catalytic center, is the most widespread and the best characterized group. Seven structures are now available. Whereas most proteins from this family are monomeric and display two LAGLIDADG motifs, a few have only one motif, and thus dimerize to cleave palindromic or pseudo-palindromic target sequences.

Although the LAGLIDADG peptide is the only conserved region among members of the family, these proteins share a very similar architecture. The catalytic core is flanked by two DNA-binding domains with a perfect two-fold symmetry for homodimers such as I-*Cre*I (Chevalier, Monnat et al. 2001), I-*Mso*I (Chevalier, Turmel et al. 2003) and I-*Ceu*I (Spiegel, Chevalier et al. 2006) and with

a pseudo symmetry for monomers such as I-*Sce*I (Moure, Gimble et al. 2003), I-*Dmo*I (Silva, Dalgaard et al. 1999) or I-*Ani*I (Bolduc, Spiegel et al. 2003). Both monomers and both domains (for monomeric proteins) contribute to the catalytic core, organized around divalent cations. Just above the catalytic core, the two LAGLIDADG peptides also play an essential role in the dimerization interface. DNA binding depends on two typical saddle-shaped αββαββα folds, sitting on the DNA major groove. Other domains can be found, for example in inteins such as PI-*Pfu*I (Ichiyanagi, Ishino et al. 2000) and PI-*Sce*I (Moure, Gimble et al. 2002), whose protein splicing domain is also involved in DNA binding.

The making of functional chimeric meganucleases, by fusing the N-terminal I-*Dmo*I domain with an I-*Cre*I monomer (Chevalier, Kortemme et al. 2002; Epinat, Arnould et al. 2003); International PCT Application WO 03/078619 (Cellectis) and WO 2004/031346 (Fred Hutchinson Cancer Research Center, Stoddard et al)) have demonstrated the plasticity of LAGLIDADG proteins.

Different groups have also used a semi-rational approach to locally alter the specificity of the I-*Cre*I (Seligman, Stephens et al. 1997; Sussman, Chadsey et al. 2004); International PCT Applications WO 2006/097784, WO 2006/097853, WO 2007/060495 and WO 2007/049156 (Cellectis); (Arnould, Chames et al. 2006; Rosen, Morrison et al. 2006; Smith, Grizot et al. 2006), I-*Sce*I (Doyon, Pattanayak et al. 2006), PI-*Sce*I (Gimble, Moure et al. 2003) and I-*Mso*I (Ashworth, Havranek et al. 2006).

In addition, hundreds of I-*Cre*I derivatives with locally altered specificity were engineered by combining the semi-rational approach and High Throughput Screening:

- Residues Q44, R68 and R70 or Q44, R68, D75 and I77 of I-*Cre*I were mutagenized and a collection of variants with altered specificity at positions ± 3 to 5 of the DNA target (5NNN DNA target) were identified by screening (International PCT Applications WO 2006/097784 and WO 2006/097853 (Cellectis); (Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

- Residues K28, N30 and Q38 or N30, Y33 and Q38 or K28, Y33, Q38 and S40 of I-*Cre*I were mutagenized and a collection of variants with altered specificity at positions ± 8 to 10 of the DNA target (10NNN DNA target) were identified by screening (Arnould, Chames et al. 2006; Smith, Grizot et al. 2006); International PCT Applications WO 2007/060495 and WO 2007/049156 (Cellectis)).

Two different variants were combined and assembled in a functional heterodimeric endonuclease able to cleave a chimeric target resulting from the fusion of two different halves of each variant DNA target sequence ((Arnould, Chames et al. 2006; Smith, Grizot et al. 2006); International PCT Applications WO 2006/097854 and WO 2007/034262).

Furthermore, residues 28 to 40 and 44 to 77 of I-*Cre*I were shown to form two partially separable functional subdomains, able to bind distinct parts of a homing endonuclease target half-

site (Smith, Grizot et al. 2006); International PCT Applications WO 2007/049095 and WO 2007/057781 (Cellectis)).

The combination of mutations from the two subdomains of I-*Cre*I within the same monomer allowed the design of novel chimeric molecules (homodimers) able to cleave a palindromic combined DNA target sequence comprising the nucleotides at positions ± 3 to 5 and ± 8 to 10 which are bound by each subdomain ((Smith, Grizot et al. 2006); International PCT Applications WO 2007/049095 and WO 2007/057781 (Cellectis)).

The method for producing meganuclease variants and the assays based on cleavage-induced recombination in mammal or yeast cells, which are used for screening variants with altered specificity are described in the International PCT Application WO 2004/067736; (Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006). These assays result in a functional LacZ reporter gene which can be monitored by standard methods.

The combination of the two former steps allows a larger combinatorial approach, involving four different subdomains. The different subdomains can be modified separately and combined to obtain an entirely redesigned meganuclease variant (heterodimer or single-chain molecule) with chosen specificity. In a first step, couples of novel meganucleases are combined in new molecules ("half-meganucleases") cleaving palindromic targets derived from the target one wants to cleave. Then, the combination of such "half-meganucleases" can result in a heterodimeric species cleaving the target of interest. The assembly of four sets of mutations into heterodimeric endonucleases cleaving a model target sequence or a sequence from different genes has been described in the following Cellectis International patent applications: XPC gene (WO2007/093918), RAG gene (WO2008/010093), HPRT gene (WO2008/059382), beta-2 microglobulin gene (WO2008/102274), Rosa26 gene (WO2008/152523), Human hemoglobin beta gene (WO2009/13622) and Human interleukin-2 receptor gamma chain gene (WO2009019614).

These variants can be used to cleave genuine chromosomal sequences and have paved the way for novel perspectives in several fields, including gene therapy.

Examples of such endonuclease include *I-Sce I, I-Chu I, I-Cre I, I-Csm I, PI-Sce I, PI-Tli I, PI-Mtu I, I-Ceu I, I-Sce II, I-Sce III, HO, PI-Civ I, PI-Ctr I, PI-Aae I, PI-Bsu I, PI-Dha I, PI-Dra I, PI-Mav I, PI-Mch I, PI-Mfu I, PI-Mfl I, PI-Mga I, PI-Mgo I, PI-Min I, PI-Mka I, PI-Mle I, PI-Mma I, PI-Msh I, PI-Msm I, PI-Mth I, PI-Mtu I, PI-Mxe I, PI-Npu I, PI-Pfu I, PI-Rma I, PI-Spb I, PI-Ssp I, PI-Fac I, PI-Mja I, PI-Pho I, PI-Tag I, PI-Thy I, PI-Tko I, PI-Tsp I, I-Msol*.

A homing endonuclease can be a LAGLIDADG endonuclease such as I-*Sce*I, I-*Cre*I, I-*Ceu*I, I-*Mso*I, and I-*Dmo*I.

Said LAGLIDADG endonuclease can be I-Sce I, a member of the family that contains two LAGLIDADG motifs and functions as a monomer, its molecular mass being approximately twice the mass of other family members like I-CreI which contains only one LAGLIDADG motif and functions as homodimers.

Endonucleases mentioned in the present application encompass both wild-type (naturally-occurring) and variant endonucleases. Endonucleases according to the invention can be a "variant" endonuclease, i.e. an endonuclease that does not naturally exist in nature and that is obtained by genetic engineering or by random mutagenesis, i.e. an engineered endonuclease. This variant endonuclease can for example be obtained by substitution of at least one residue in the amino acid sequence of a wild-type, naturally-occurring, endonuclease with a different amino acid. Said substitution(s) can for example be introduced by site-directed mutagenesis and/or by random mutagenesis. In the frame of the present invention, such variant endonucleases remain functional, i.e. they retain the capacity of recognizing (binding function) and optionally specifically cleaving a target sequence to initiate gene targeting process.

The variant endonuclease according to the invention cleaves a target sequence that is different from the target sequence of the corresponding wild-type endonuclease. Methods for obtaining such variant endonucleases with novel specificities are well-known in the art.

Endonucleases variants may be homodimers (meganuclease comprising two identical monomers) or heterodimers (meganuclease comprising two non-identical monomers). It is understood that the scope of the present invention also encompasses endonuclease variants *per se*, including heterodimers (WO2006097854), obligate heterodimers (WO2008093249) and single chain meganucleases (WO03078619 and WO2009095793) as non limiting examples, able to cleave one target of interest in a polynucleotidic sequence or in a genome. The invention also encompasses hybrid variant *per se* composed of two monomers from different origins (WO03078619).

Endonucleases with novel specificities can be used in the method according to the present invention for gene targeting and thereby integrating a transgene of interest into a genome at a predetermined location.

- by "parent meganuclease" it is intended to mean a wild type meganuclease or a variant of such a wild type meganuclease with identical properties or alternatively a meganuclease with some altered characteristics in comparison to a wild type version of the same meganuclease. This expression can also be transposed to an endonuclease, a rare-cutting endonuclease, a chimeric rare-cutting endonuclease, a TALEN or a compact TALEN and derivatives.

- By " delivery vector" or " delivery vectors" is intended any delivery vector which can be used in the present invention to put into cell contact (i.e "contacting") or deliver inside cells or

subcellular compartments agents/chemicals and molecules (proteins or nucleic acids) needed in the present invention. It includes, but is not limited to liposomal delivery vectors, viral delivery vectors, drug delivery vectors, chemical carriers, polymeric carriers, lipoplexes, polyplexes, dendrimers, microbubbles (ultrasound contrast agents), nanoparticles, emulsions or other appropriate transfer vectors. These delivery vectors allow delivery of molecules, chemicals, macromolecules (genes, proteins), or other vectors such as plasmids, peptides developed by Diatos. In these cases, delivery vectors are molecule carriers. By "delivery vector" or "delivery vectors" is also intended delivery methods to perform transfection.

- The terms "vector" or "vectors" refer to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. A "vector" in the present invention includes, but is not limited to, a viral vector, a plasmid, a RNA vector or a linear or circular DNA or RNA molecule which may consists of a chromosomal, non chromosomal, semi-synthetic or synthetic nucleic acids. Preferred vectors are those capable of autonomous replication (episomal vector) and/or expression of nucleic acids to which they are linked (expression vectors). Large numbers of suitable vectors are known to those of skill in the art and commercially available.

Viral vectors include retrovirus, adenovirus, parvovirus (e. g. adenoassociated viruses), coronavirus, negative strand RNA viruses such as orthomyxovirus (e. g., influenza virus), rhabdovirus (e. g., rabies and vesicular stomatitis virus), paramyxovirus (e. g. measles and Sendai), positive strand RNA viruses such as picornavirus and alphavirus, and double-stranded DNA viruses including adenovirus, herpesvirus (e. g., Herpes Simplex virus types 1 and 2, Epstein-Barr virus, cytomegalovirus), and poxvirus (e. g., vaccinia, fowlpox and canarypox). Other viruses include Norwalk virus, togavirus, flavivirus, reoviruses, papovavirus, hepadnavirus, and hepatitis virus, for example. Examples of retroviruses include: avian leukosis-sarcoma, mammalian C-type, B-type viruses, D type viruses, HTLV-BLV group, lentivirus, spumavirus (Coffin, J. M., Retroviridae: The viruses and their replication, In Fundamental Virology, Third Edition, B. N. Fields, et al., Eds., Lippincott-Raven Publishers, Philadelphia, 1996).

- By "lentiviral vector" is meant HIV-Based lentiviral vectors that are very promising for gene delivery because of their relatively large packaging capacity, reduced immunogenicity and their ability to stably transduce with high efficiency a large range of different cell types. Lentiviral vectors are usually generated following transient transfection of three (packaging, envelope and transfer) or more plasmids into producer cells. Like HIV, lentiviral vectors enter the target cell through the interaction of viral surface glycoproteins with receptors on the cell surface. On entry, the viral RNA undergoes reverse transcription, which is mediated by the viral reverse transcriptase

complex. The product of reverse transcription is a double-stranded linear viral DNA, which is the substrate for viral integration in the DNA of infected cells.

- By "integrative lentiviral vectors (or LV)", is meant such vectors as non limiting example, that are able to integrate the genome of a target cell.
- At the opposite by "non integrative lentiviral vectors (or NILV)" is meant efficient gene delivery vectors that do not integrate the genome of a target cell through the action of the virus integrase.

One type of preferred vector is an episome, i.e., a nucleic acid capable of extra-chromosomal replication. Preferred vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as "expression vectors. A vector according to the present invention comprises, but is not limited to, a YAC (yeast artificial chromosome), a BAC (bacterial artificial), a baculovirus vector, a phage, a phagemid, a cosmid, a viral vector, a plasmid, a RNA vector or a linear or circular DNA or RNA molecule which may consist of chromosomal, non chromosomal, semi-synthetic or synthetic DNA. In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer generally to circular double stranded DNA loops which, in their vector form are not bound to the chromosome. Large numbers of suitable vectors are known to those of skill in the art. Vectors can comprise selectable markers, for example: neomycin phosphotransferase, histidinol dehydrogenase, dihydrofolate reductase, hygromycin phosphotransferase, herpes simplex virus thymidine kinase, adenosine deaminase, glutamine synthetase, and hypoxanthine-guanine phosphoribosyl transferase for eukaryotic cell culture; TRP1 for *S. cerevisiae*; tetracyclin, rifampicin or ampicillin resistance in *E. coli*. Preferably said vectors are expression vectors, wherein a sequence encoding a polypeptide of interest is placed under control of appropriate transcriptional and translational control elements to permit production or synthesis of said polypeptide. Therefore, said polynucleotide is comprised in an expression cassette. More particularly, the vector comprises a replication origin, a promoter operatively linked to said encoding polynucleotide, a ribosome binding site, a RNA-splicing site (when genomic DNA is used), a polyadenylation site and a transcription termination site. It also can comprise an enhancer or silencer elements. Selection of the promoter will depend upon the cell in which the polypeptide is expressed. Suitable promoters include tissue specific and/or inducible promoters. Examples of inducible promoters are: eukaryotic metallothionine promoter which is induced by increased levels of heavy metals, prokaryotic lacZ promoter which is induced in response to isopropyl-β-D-thiogalacto-pyranoside (IPTG) and eukaryotic heat shock promoter which is induced by increased temperature. Examples of tissue specific promoters are skeletal muscle creatine kinase,

prostate-specific antigen (PSA), α-antitrypsin protease, human surfactant (SP) A and B proteins, β-casein and acidic whey protein genes.

- Inducible promoters may be induced by pathogens or stress, more preferably by stress like cold, heat, UV light, or high ionic concentrations (reviewed in Potenza C et al. 2004, In vitro Cell Dev Biol 40:1-22). Inducible promoter may be induced by chemicals (reviewed in (Moore, Samalova et al. 2006); (Padidam 2003); (Wang, Zhou et al. 2003); (Zuo and Chua 2000).

Delivery vectors and vectors can be associated or combined with any cellular permeabilization techniques such as sonoporation or electroporation or derivatives of these techniques.

- By cell or cells is intended any prokaryotic or eukaryotic living cells, cell lines derived from these organisms for *in vitro* cultures, primary cells from animal or plant origin.

- By "primary cell" or "primary cells" are intended cells taken directly from living tissue (i.e. biopsy material) and established for growth in vitro, that have undergone very few population doublings and are therefore more representative of the main functional components and characteristics of tissues from which they are derived from, in comparison to continuous tumorigenic or artificially immortalized cell lines. These cells thus represent a more valuable model to the in vivo state they refer to.

- In the frame of the present invention, "eukaryotic cells" refer to a fungal, plant or animal cell or a cell line derived from the organisms listed below and established for in vitro culture. More preferably, the fungus is of the genus Aspergillus, Penicillium, Acremonium, Trichoderma, Chrysoporium, Mortierella, Kluyveromyces or Pichia; More preferably, the fungus is of the species Aspergillus niger, Aspergillus nidulans, Aspergillus oryzae, Aspergillus terreus, Penicillium chrysogenum, Penicillium citrinum, Acremonium Chrysogenum, Trichoderma reesei, Mortierella alpine, Chrysosporium lucknowense, Kluyveromyces lactis, Pichia pastoris or Pichia ciferrii.

More preferably the plant is of the genus Arabidospis, Nicotiana, Solanum, lactuca, Brassica, Oryza, Asparagus, Pisum, Medicago, Zea, Hordeum, Secale, Triticum, Capsicum, Cucumis, Cucurbita, Citrullis, Citrus, Sorghum; More preferably, the plant is of the species Arabidospis thaliana, Nicotiana tabaccum, Solanum lycopersicum, Solanum tuberosum, Solanum melongena, Solanum esculentum, Lactuca saliva, Brassica napus, Brassica oleracea, Brassica rapa, Oryza glaberrima, Oryza sativa, Asparagus officinalis, Pisum sativum, Medicago sativa, zea mays, Hordeum vulgare, Secale cereal, Triticum aestivum, Triticum durum, Capsicum sativus, Cucurbita pepo, Citrullus lanatus, Cucumis melo, Citrus aurantifolia, Citrus maxima, Citrus medica, Citrus reticulata.

More preferably the animal cell is of the genus Homo, Rattus, Mus, Sus, Bos, Danio, Canis, Felis, Equus, Salmo, Oncorhynchus, Gallus, Meleagris, Drosophila, Caenorhabditis; more preferably,

the animal cell is of the species Homo sapiens, Rattus norvegicus, Mus musculus, Sus scrofa, Bos taurus, Danio rerio, Canis lupus, Felis catus, Equus caballus, Salmo salar, Oncorhynchus mykiss, Gallus gallus, Meleagris gallopavo, Drosophila melanogaster, Caenorhabditis elegans.

In the present invention, the cell can be a plant cell, a mammalian cell, a fish cell, an insect cell or cell lines derived from these organisms for *in vitro* cultures or primary cells taken directly from living tissue and established for *in vitro* culture. As non-limiting examples, cell can be protoplasts obtained from plant organisms listed above. As non limiting examples cell lines can be selected from the group consisting of CHO-K1 cells; HEK293 cells; Caco2 cells; U2-OS cells; NIH 3T3 cells; NSO cells; SP2 cells; CHO-S cells; DG44 cells; K-562 cells, U-937 cells; MRC5 cells; IMR90 cells; Jurkat cells; HepG2 cells; HeLa cells; HT-1080 cells; HCT-116 cells; Hu-h7 cells; Huvec cells; Molt 4 cells.

All these cell lines can be modified by the method of the present invention to provide cell line models to produce, express, quantify, detect, study a gene or a protein of interest; these models can also be used to screen biologically active molecules of interest in research and production and various fields such as chemical, biofuels, therapeutics and agronomy as non-limiting examples. Adoptive immunotherapy using genetically engineered T cells is a promising approach for the treatment of malignancies and infectious diseases. Most current approaches rely on gene transfer by random integration of an appropriate T Cell Receptor (TCR) or Chimeric Antigen Receptor (CAR). Targeted approach using rare-cutting endonucleases is an efficient and safe alternative method to transfer genes into T cells and generate genetically engineered T cells.

- by "homologous" is intended a sequence with enough identity to another one to lead to homologous recombination between sequences, more particularly having at least 95 % identity, preferably 97 % identity and more preferably 99 %.

- "identity" refers to sequence identity between two nucleic acid molecules or polypeptides. Identity can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base, then the molecules are identical at that position. A degree of similarity or identity between nucleic acid or amino acid sequences is a function of the number of identical or matching nucleotides at positions shared by the nucleic acid sequences. Various alignment algorithms and/or programs may be used to calculate the identity between two sequences, including FASTA, or BLAST which are available as a part of the GCG sequence analysis package (University of Wisconsin, Madison, Wis.), and can be used with, e.g., default setting.

- by "mutation" is intended the substitution, deletion, insertion of one or more nucleotides/amino acids in a polynucleotide (cDNA, gene) or a polypeptide sequence. Said mutation

can affect the coding sequence of a gene or its regulatory sequence. It may also affect the structure of the genomic sequence or the structure/stability of the encoded mRNA.

- In the frame of the present invention, the expression "double-strand break-induced mutagenesis" (DSB-induced mutagenesis) refers to a mutagenesis event consecutive to an NHEJ event following an endonuclease-induced DSB, leading to insertion/deletion at the cleavage site of an endonuclease.

- By "gene" is meant the basic unit of heredity, consisting of a segment of DNA arranged in a linear manner along a chromosome, which codes for a specific protein or segment of protein. A gene typically includes a promoter, a 5' untranslated region, one or more coding sequences (exons), optionally introns, a 3' untranslated region. The gene may further comprise a terminator, enhancers and/or silencers.

- As used herein, the term "transgene" refers to a sequence encoding a polypeptide. Preferably, the polypeptide encoded by the transgene is either not expressed, or expressed but not biologically active, in the cell, tissue or individual in which the transgene is inserted. Most preferably, the transgene encodes a therapeutic polypeptide useful for the treatment of an individual.

- The term "gene of interest" or "GOI" refers to any nucleotide sequence encoding a known or putative gene product.

- As used herein, the term "locus" is the specific physical location of a DNA sequence (e.g. of a gene) on a chromosome. The term "locus" usually refers to the specific physical location of an endonuclease's target sequence on a chromosome. Such a locus, which comprises a target sequence that is recognized and cleaved by an endonuclease according to the invention, is referred to as "locus according to the invention". Also, the expression "genomic locus of interest" is used to qualify a nucleic acid sequence in a genome that can be a putative target for a double-strand break according to the invention. It is understood that the considered genomic locus of interest of the present invention can not only qualify a nucleic acid sequence that exists in the main body of genetic material (i.e. in a chromosome) of a cell but also a portion of genetic material that can exist independently to said main body of genetic material such as plasmids, episomes, virus, transposons or in organelles such as mitochondria or chloroplasts as non-limiting examples.

- By the expression "loss of genetic information" is understood the elimination or addition of at least one given DNA fragment (at least one nucleotide) or sequence, bordering the recognition sites of the endonucleases, chimeric rare-cutting endonucleases, compact TALEN or enhanced

compact TALEN of the present invention or the intervening sequence between at least two processing sites of the endonucleases, chimeric rare-cutting endonucleases, compact TALEN or enhanced compact TALEN of the present invention and leading to a change of the original sequence around said endonuclease-cutting sites, chimeric rare-cutting endonuclease-cutting sites, compact TALEN or enhanced compact TALEN recognition DNA binding site within the genomic locus of interest. This loss of genetic information can be, as a non-limiting example, the elimination of an intervening sequence between two endonuclease-cutting sitesor between two processing sites of a compact TALEN or enhanced compact TALEN. As another non-limiting example, this loss of genetic information can also be an excision of a single-strand of DNA spanning the binding region of a compact TALEN or an enhanced compact TALEN according to the present invention

- By "scarless re-ligation" or "scarless religation" is intended the perfect re-ligation event, without loss of genetic information (no insertion/deletion events) of the DNA broken ends through NHEJ process after the creation of a double-strand break event.

- By "Imprecise NHEJ" is intended the re-ligation of nucleic acid ends generated by a DSB, with insertions or deletions of nucleotides. Imprecise NHEJ is an outcome and not a repair pathway and can result from different NHEJ pathways (Ku dependent or Ku independent as non-limiting examples).

- By "fusion protein" is intended the result of a well-known process in the art consisting in the joining of two or more genes which originally encode for separate proteins or part of them, the translation of said "fusion gene" resulting in a single polypeptide with functional properties derived from each of the original proteins.

- By "chimeric rare-cutting endonuclease" is meant any fusion protein comprising a rare-cutting endonuclease. Said rare-cutting endonuclease might be at the N- terminal part of said chimeric rare-cutting endonuclease; at the opposite, said rare-cutting endonuclease might be at the C- terminal part of said chimeric rare-cutting endonuclease. A "chimeric rare-cutting endonuclease" according to the present invention which comprises two catalytic domains can be described as "bi-functional" or as "bi-functional meganuclease". A "chimeric rare-cutting endonuclease" according to the present invention which comprises more than two catalytic domains can be described as "multi-functional" or as "multi-functional meganuclease". As non-limiting examples, chimeric rare-cutting endonucleases according to the present invention can be a fusion protein between a rare-cutting endonuclease and one catalytic domain; chimeric rare-cutting endonucleases according to the present invention can also be a fusion protein between a rare-cutting endonuclease and two catalytic domains. As mentioned previously, the rare-cutting endonuclease part of chimeric rare-cutting endonucleases according to the present invention can be a meganuclease comprising two

identical monomers, two non-identical monomers, or a single chain meganuclease. The rare-cutting endonuclease part of chimeric rare-cutting endonucleases according to the present invention can also be the DNA-binding domain of an inactive rare-cutting endonuclease. In other non-limiting examples, chimeric rare-cutting endonucleases according to the present invention can be derived from a TALE-nuclease (TALEN), i. e. a fusion between a DNA-binding domain derived from a Transcription Activator Like Effector (TALE) and one or two catalytic domains. In other non-limiting examples, a subclass of chimeric rare-cutting endonucleases according to the present invention can be a "compact TALE-nuclease" (cTALEN), i. e. a fusion between an engineered core TALE scaffold comprising at least a Repeat Variable Dipeptide regions domain and at least one catalytic domain, said fusion protein constituting a compact TALEN active entity that does not require dimerization for DNA processing activity. Said catalytic domain can be an endonuclease as listed in table 2 as non-limiting examples; said catalytic domain can be a frequent-cutting endonuclease such as a restriction enzyme selected from the group consisting of MmeI, R-HinPII, R.MspI, R.MvaI, Nb.BsrDI, BsrDI A, Nt.BspD6I, ss.BspD6I, R.PleI, MlyI and AlwI as non-limiting restriction enzymes examples listed in table 2.

- By "enhancer domain(s)" or "enhancer(s)" are meant protein domains that provide functional and/or structural support to a protein scaffold, a compact TALEN as a non-limiting example, therefore allowing an enhancement in global DNA processing efficiency of the resulting fusion protein, i.e an enhanced compact TALEN, relative to the DNA processing efficiency of the starting compact TALEN. A particular domain is an enhancer domain when it provides at least a 5% enhancement in efficiency of the starting scaffold, more preferably 10 %, again more preferably 15 %, again more preferably 20 %, again more preferably 25 %, again more preferably 50%, again more preferably greater than 50%. Non-limiting examples of such enhancer domains are given in Tables 1 and 2. By "auxiliary enhancer domains" or "auxiliary enhancers" or "auxiliary domains" are meant protein domains acting in trans with a compact TALEN or an enhanced compact TALEN to provide an additional function that is not essential for said basic compact TALEN activity or said enhanced compact TALEN activity. When such auxiliary enhancers are used, compact TALEN or enhanced compact TALEN are converted to "trans TALEN", respectively trans compact TALEN and trans enhanced compact TALEN.

- By "catalytic domain" is intended the protein domain or module of an enzyme containing the active site of said enzyme; by active site is intended the part of said enzyme at which catalysis of the substrate occurs. Enzymes, but also their catalytic domains, are classified and named according to the reaction they catalyze. The Enzyme Commission number (EC number) is a numerical

classification scheme for enzymes, based on the chemical reactions they catalyze (http://www.chem.qmul.ac.uk/iubmb/enzyme/). In the scope of the present invention, any catalytic domain can be fused to an engineered core TALE scaffold to generate a compact TALEN active entity with a DNA processing efficiency provided by at least said catalytic domain activity. Said catalytic domain can provide a nuclease activity (endonuclease or exonuclease activities, cleavase or nickase), a polymerase activity, a kinase activity, a phosphatase activity, a methylase activity, a topoisomerase activity, an integrase activity, a transposase activity or a ligase activity as non-limiting examples. Non-limiting examples of such catalytic domains are given in Tables 1 and 2. In a preferred embodiment of the present invention, said catalytic domain can be considered as an enhancer domain. If catalytically active, said enhancer domain can provide functional and/or structural support to the compact TALEN scaffold leading to an enhanced compact TALEN when fused to it. If catalytically inactive, said enhancer domain provides structural support to compact TALEN scaffold leading to an enhanced compact TALEN when fused to it. It can be envisioned from the present invention to fuse catalytic domains according to the present invention to one part of a classical TALEN in order to give these classical TALENs new catalytical properties provided by at least said catalytic domain activity or to improve their DNA processing efficiency.

- By "nuclease catalytic domain" is intended the protein domain comprising the active site of an endonuclease or an exonuclease enzyme. Such nuclease catalytic domain can be, for instance, a "cleavase domain" or a "nickase domain". By "cleavase domain" is intended a protein domain whose catalytic activity generates a Double Strand Break (DSB) in a DNA target. By "nickase domain" is intended a protein domain whose catalytic activity generates a single strand break in a DNA target sequence. Non-limiting examples of such catalytic domains are given in Tables 1 and 2 with a GenBank or NCBI or UniProtKB/Swiss-Prot number as a reference.

- By a "TALE-nuclease" (TALEN) or a "classical TALEN" is intended a fusion protein consisting of a DNA-binding domain derived from a Transcription Activator Like Effector (TALE) and one FokI catalytic domain, that need to dimerize to form an active entity able to cleave a DNA target sequence.

- By "compact TALE-nuclease" (cTALEN) is intended a general designation according to the present invention for a fusion protein between an engineered core TALE scaffold comprising at least one Repeat Variable Dipeptides domain and at least one catalytic domain, said fusion protein constituting a compact TALEN (or cTALEN) active entity and not requiring dimerization for DNA processing activity. Compact TALENs are designed to alleviate the need for multiple independent protein moieties when targeting a DNA cleavage event. Importantly, the requisite "spacer" region and dual target sites essential for the function of current TALENs are unnecessary. In other words,

the compact TALEN according to the present invention is an active entity unit able, by itself, to target only one specific single double-stranded DNA target sequence of interest through one DNA binding domain and to process DNA nearby said single double-stranded DNA target sequence of interest. In addition, since the catalytic domain does not require specific DNA contact, there are no restrictions on regions surrounding the core TALE DNA binding domain. In the scope of the present invention, it can be also envisioned some sequence preference in the catalytic domain. When a cTALEN comprises only one catalytic domain, cTALEN can be qualified as a "basic cTALEN" or "cTALEN". When a cTALEN further comprises at least one "enhancer domain", cTALEN can be qualified as an enhanced cTALEN or an "eTALEN". A cTALEN or an eTALEN that comprise at least one cleavase catalytic domain and one nickase catalytic domain or at least two cleavase catalytic domains can be specifically qualified as a dual-cleavage cTALEN or a "dcTALEN". A cTALEN or an eTALEN acting with an auxiliary domain in trans is qualified as a trans compact TALEN or a trans enhanced compact TALEN, both being "trans TALEN".

The above written description of the invention provides a manner and process of making and using it such that any person skilled in this art is enabled to make and use the same, this enablement being provided in particular for the subject matter of the appended claims, which make up a part of the original description.

As used above, the phrases "selected from the group consisting of," "chosen from," and the like include mixtures of the specified materials.

Where a numerical limit or range is stated herein, the endpoints are included. Also, all values and subranges within a numerical limit or range are specifically included as if explicitly written out.

The above description is presented to enable a person skilled in the art to make and use the invention, and is provided in the context of a particular application and its requirements. Various modifications to the preferred embodiments will be readily apparent to those skilled in the art, and the generic principles defined herein may be applied to other embodiments and applications without departing from the spirit and scope of the invention. Thus, this invention is not intended to be limited to the embodiments shown, but is to be accorded the widest scope consistent with the principles and features disclosed herein.

Having generally described this invention, a further understanding can be obtained by reference to certain specific examples, which are provided herein for purposes of illustration only, and are not intended to be limiting unless otherwise specified.

Examples

Example 1

The wild-type I-CreI meganuclease (SEQ ID NO: 106) was chosen as the parent scaffold on which to fuse the catalytic domain of I-TevI (SEQ ID NO: 107). Wild-type I-TevI functions as a monomeric cleavase of the GIY-YIG family to generate a staggered double-strand break in its target DNA . Guided by biochemical and structural data, variable length constructs were designed from the N-terminal region of I-TevI that encompass the entire catalytic domain and deletion-intolerant region of its linker (SEQ ID NO: 109 to SEQ ID NO: 114). In all but one case, fragments were fused to the N-terminus of I-CreI with an intervening 5-residue polypeptide linker (-QGPSG-; SEQ ID NO: 103). The linker-less fusion construct naturally contained residues (-LGPDGRKA-; SEQ ID NO: 104) similar to those in the artificial linker. As I-CreI is a homodimer, all fusion constructs contain three catalytic centers (Figure 4, where "catalytic domain" = cleavase): the natural I-CreI active site at the interface of the dimer and one I-TevI active site per monomer.

The activity of each "tri-functional" meganuclease was assessed using our yeast assay previously described in International PCT Applications WO 2004/067736 and in (Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006). All constructs were able to cleave the C1221 target DNA with an activity comparable to that of wild-type I-CreI (Table 4).

To validate the activity of the I-TevI catalytic domain independent of the I-CreI catalytic core, D20N point mutants were made to inactivate the I-CreI scaffold [SEQ ID NO: 108, SEQ ID NO: 115 to SEQ ID NO: 120; Chevalier, Sussman et al. 2004)]. Tests in our yeast assays showed no visible activity from the inactivated I-CreI (D20N) mutant protein alone (Table 4). However, cleavage activity could be observed for fusions having the I-TevI catalytic domain (Table 4).

Protein Construct	Relative Activity in Yeast Assay (37°C)	
	C1221 Target	Tev Target
I-CreI	++++	-
I-TevI	-	++++
I-CreI_N20	-	-
hTevCre_D01	++++	-
hTevCre_D02	++++	-

hTevCre_D03	++++	-
hTevCre_D04	++++	-
hTevCre_D05	++++	-
hTevCre_D06	++++	-
hTevCre_D01_N20	++	-
hTevCre_D02_N20	++	-
hTevCre_D03_N20	++	-
hTevCre_D04_N20	++	-
hTevCre_D05_N20	-	-
hTevCre_D06_N20	-	-

Table 4: Activity in Yeast assay for I-TevI/I-CreI fusions. The relative activity of wild-type and fusion proteins on the two parent protein targets (C1221 for I-CreI and Tev for I-TevI) is shown. Maximal activity (++++) is seen with each given protein on its native DNA target. I-CreI_N20 is an inactive variant of the wild-type I-CreI scaffold. In all other cases, activity is only detected on the C1221 target since DNA recognition is driven by the I-CreI scaffold. The "N20" fusion variants illustrate cleavage activity due to the I-TevI catalytic domain.
Relative activity is scaled as: -, no activity detectable; +, <25% activity; ++, 25% to <50% activity; +++, 50% to <75% activity; ++++, 75% to 100% activity.

Example 2

Protein-fusion scaffolds were designed based on a truncated form of I-CreI (SEQ ID NO: 106, I-CreI_X: SEQ ID NO: 121) and three different linker polypeptides (NFS1 = SEQ ID NO: 98; NFS2 = SEQ ID NO: 99; CFS1 = SEQ ID NO: 100) fused to either the N- or C-terminus of the protein. Structure models were generated in all cases, with the goal of designing a "baseline" fusion linker that would traverse the I-CreI parent scaffold surface with little to no effect on its DNA binding or cleavage activities. For the two N-terminal fusion scaffolds, the polypeptide spanning residues 2 to 153 of I-CreI was used, with a K82A mutation to allow for linker placement. The C-terminal fusion scaffold contains residues 2 to 155 of wild-type I-CreI. For both fusion scaffold types, the "free" end of the linker (i.e. onto which a polypeptide can be linked) is designed to be proximal to the DNA, as determined from models built using the I-CreI/DNA complex structures as a starting point (PDB id: 1g9z). The two I-CreI N-terminal fusion scaffolds (I-CreI_NFS1 = SEQ ID NO: 122 and I-CreI_NFS2 = SEQ ID NO: 123) and the single C-terminal fusion scaffold (I-CreI_CFS1 = SEQ ID NO: 124) were tested in our yeast assay (see Example 1) and found to have activity similar to that of wild-type I-CreI (Table 5).

Protein Construct	**Relative Activity in Yeast Assay (37°C)**
	C1221 Target
I-CreI	++++
I-CreI_X	++++

I-CreI_NFS1	++++
I-CreI_NFS2	++++
I-CreI_CFS1	++++
I-CreI_NFS1_N20	-
I-CreI_NFS2_N20	-
I-CreI_CFS1_N20	-
hColE7Cre_D0101	++++
hColE7Cre_D0102	++++
hCreColE7_D0101	++++
hColE7Cre_D0101_N20	+++
hColE7Cre_D0102_N20	+++
hCreColE7_D0101_N20	++

Table 5: Activity in Yeast assay for ColE7/I-CreI fusions. The relative activity of wild-type and fusion proteins on theC1221 target is shown. I-CreI_X represents a truncated version of I-CreI based on the crystal structure and was used as the foundation for the fusion scaffolds (I-CreI_NFS1, I-CreI_NFS2 and I-CreI_CFS1). "N20" constructs are inactive variants of the respective I-CreI-based scaffolds. Activity is detected in all cases wherein the I-CreI scaffold is active or when DNA catalysis is provided by the ColE7 domain.
Relative activity is scaled as: -, no activity detectable; +, <25% activity; ++, 25% to <50% activity; +++, 50% to <75% activity; ++++, 75% to 100% activity.

Colicin E7 is a non-specific nuclease of the HNH family able to process single- and double-stranded DNA (Hsia, Chak et al. 2004). Guided by biochemical and structural data, the region of ColE7 that encompasses the entire catalytic domain (SEQ ID NO: 140; (Hsia, Chak et al. 2004) was selected. This ColE7 domain was fused to the N-terminus of either I-CreI_NFS1 (SEQ ID NO: 122) or I-CreI_NFS2 (SEQ ID NO: 123) to create hColE7Cre_D0101 (SEQ ID NO: 128) or hColE7Cre_D0102 (SEQ ID NO: 129), respectively. In addition, a C-terminal fusion construct, hCreColE7_D0101 (SEQ ID NO: 130), was generated using I-CreI_CFS1 (SEQ ID NO: 124). As I-CreI is a homodimer, all fusion constructs contain three catalytic centers (Figure 4, where "catalytic domain" = cleavase): the natural I-CreI active site at the interface of the dimer and one ColE7 active site per monomer.

The activity of each "tri-functional" meganuclease was assessed using our yeast assay (see Example 1). All constructs were able to cleave the C1221 target DNA with an activity comparable to that of wild-type I-CreI (Table 5).

To validate the activity of the ColE7 catalytic domain independent of the I-CreI catalytic core, D20N point mutants were made to inactivate the I-CreI scaffold (SEQ ID NO: 125, SEQ ID NO: 126, SEQ ID NO: 127, SEQ ID NO: 131, SEQ ID NO: 132, SEQ ID NO: 133; (Chevalier, Sussman et al. 2004)). Tests in our yeast assays showed no visible activity from the inactivated I-CreI (D20N) mutant proteins alone

(Table 5). However, cleavage activity could be observed for fusions having the ColE7 catalytic domain (Table 5).

Example 3

Two core TALE scaffolds are generated onto which (a) different sets of RVD domains could be inserted to change DNA binding specificity, and; (b) a selection of catalytic domains could be attached, N- or C-terminal, to effect DNA cleavage (or nicking). The core scaffolds (sT1: SEQ ID NO: 134 and sT2: SEQ ID NO: 135) differ in the N- and C-terminal regions, where sT2 is a truncated variant lacking 152 amino acid residues from the N-terminus (Szurek, Rossier et al. 2002) and the last 220 residues from the C-terminus compared to sT1. In sT1, the C-terminal region is a truncation with respect to wild-type TALE domains, ending at a fortuitously defined restriction site (BamHI) in the DNA coding sequence.

Using the two core scaffolds, four "baseline" TALE DNA binding proteins (bT1-Avr = SEQ ID NO: 136, bT2-Avr = SEQ ID NO: 137, bT1-Pth = SEQ ID NO 138 and bT2-Pth = SEQ ID NO 139) are generated by insertion of the corresponding set of repeat domains that recognize the naturally occurring asymmetric sequences AvrBs3 (19 bp) and PthXo1 (25 bp) (Figure 3). Example protein sequences of the baseline scaffolds are listed in SEQ ID NO: 136 to SEQ ID NO: 139. As is, these scaffolds can be tested in vitro for DNA binding ability on targets having only a single recognition sequence. For comparison with existing TALENs, the catalytic domain of the FokI nuclease (SEQ ID NO: 368 and particularly residues P381 to F583 as non-limiting example) can be fused to either the N- or C-terminus of the baseline scaffolds. Effective cleavage using these controls requires target site DNAs that contain two TALE binding sequences.

In addition to verifying activity using naturally occurring sequences, five artificial RVD constructs recognizing relevant sequences were generated (Figure 3): RagT2-R, NptIIT5-L, NptIIT5-R, NptIIT6-L, NptIIT6-R. Example protein sequences of the insert RVDs are listed in SEQ ID NO: 253 to SEQ ID NO: 257. Artificial RVD sequences are used as noted above within the sT1 or sT2 scaffold to generate the desired targeted compact TALENs.

Basic compact TALENs (cTALENs) are generated via fusion of catalytic domains to either the N- or C-terminus of the baseline scaffolds (Figure 5, A or B, respectively). A non-exhaustive list of catalytic domains amenable to fusion with TALE DNA binding domains is presented in Table 2. A non-

exhaustive list of linkers that can be used is presented in Table 3. It is notable that linker design can depend on the nature of the catalytic domain attached and its given application. It can also be anticipated that specially engineered linkers can be constructed to better control or regulate the activity of either or both domains. Examples 5, 6 and 7 below discuss additional and alternative methods in which linkers can be defined. All cTALEN designs are assessed using our yeast assay (see Example 1) and provide detectable activity comparable to existing engineered meganucleases.

Example 3a: TALE::TevI compact TALEN

The catalytic domain of I-TevI (SEQ ID NO: 20), a member of the GIY-YIG endonuclease family, was fused to a TALE-derived scaffold (composed of a N-terminal domain, a central core composed of RVDs and a C-terminal domain) to create a new class of cTALEN (TALE::TevI). To distinguish the orientation (N-terminal vs. C-terminal) of the catalytic domain (CD) fusions, construct names are written as either CD::TALE-RVD (catalytic domain is fused N-terminal to the TALE domain) or TALE-RVD::CD (catalytic domain is fused C-terminal to the TALE domain), where "-RVD" optionally designates the sequence recognized by the TALE domain and "CD" is the catalytic domain type. Herein, we describe novel TALE::TevI constructions that target AvrBs3 sequence for example, thus named TALE-AvrBs3::TevI.

Activity of TALE::TevI in yeast

A core TALE scaffold, sT2 (SEQ ID NO: 135), was selected onto which (a) different sets of RVD domains could be inserted to change DNA binding specificity, and; (b) a selection of I-TevI-derived catalytic domains could be attached, N- or C-terminal, to effect DNA cleavage (or nicking). The previously mentioned sT2 truncated scaffold was generated by the PCR from a full-length core TALEN scaffold template (pCLS7183, SEQ ID NO: 141) using primers CMP_G061 (SEQ ID NO: 142) and CMP_G065 (SEQ ID NO: 143) and was cloned into vector pCLS7865 (SEQ ID NO: 144) to generate pCLS7865-cTAL11_CFS1 (pCLS9009, SEQ ID NO: 145), where CFS1 designates the amino acid sequence -GSSG- (with underlying restriction sites BamHI and Kpn2I in the coding DNA to facilitate cloning). Three variants of the I-TevI (SEQ ID NO: 20) catalytic domain were amplified by the PCR on templates TevCreD01 [SEQ ID NO: 109 protein in plasmid pCLS6614 (SEQ ID NO: 146)] using the primer pair CMP_G069 (SEQ ID NO: 147) and CMP_G070 (SEQ ID NO: 148), TevCreD02 [SEQ ID NO: 110 protein in plasmid pCLS6615 (SEQ ID NO: 203)] using the primer pair CMP_G069 (SEQ ID NO: 147) and CMP_G071 (SEQ ID NO: 149) or TevCreD05 [SEQ ID NO: 113 protein in plasmid pCLS6618

(SEQ ID NO: 258)] using the primer pair CMP_G069 (SEQ ID NO: 147) and CMP_G115 (SEQ ID NO: 259) and subcloned into the pCLS9009 backbone by restriction and ligation using BamHI and EagI restriction sites, yielding pCLS7865-cT11_TevD01 (pCLS9010, SEQ ID NO: 150), pCLS7865-cT11_TevD02 (pCLS9011, SEQ ID NO: 151) and pCLS7865-cT11_TevD05 (pCLS15775, SEQ ID NO: 260), respectively. All fusions contain the dipeptide -GS- linking the TALE-derived DNA binding domain and I-TevI-derived catalytic domain.

The DNA sequence coding for the RVDs to target the AvrBs3 site (SEQ ID NO: 152) was subcloned into both plasmids pCLS9010 (SEQ ID NO: 150, encoding the protein of SEQ ID NO: 420), pCLS9011 (SEQ ID NO: 151, encoding the protein of SEQ ID NO: 421) and pCLS15775 (SEQ ID NO: 260, encoding the protein of SEQ ID NO: 422) using Type IIS restriction enzymes BsmBI for the receiving plasmid and BbvI and SfaNI for the inserted RVD sequence to create the subsequent TALE-AvrBs3::TevI constructs cT11Avr_TevD01 (pCLS9012, SEQ ID NO: 218, encoding the protein of SEQ ID NO: 423), cT11Avr_TevD02 (pCLS9013, SEQ ID NO: 153, encoding the protein of SEQ ID NO: 424) and cT11Avr_TevD05 (pCLS15776, SEQ ID NO: 261, encoding the protein of SEQ ID NO: 425), respectively. These TALE-AvrBs3::TevI constructs were sequenced and the insert transferred to additional vectors as needed (see below).

The final TALE-AvrBs3::TevI yeast expression plasmids, pCLS8523 (SEQ ID NO: 154), pCLS8524 (SEQ ID NO: 155) and pCLS12092 (SEQ ID NO: 262), were prepared by yeast *in vivo* cloning using plasmids pCLS9012, pCLS9013 and pCLS15776, respectively. To generate an intact coding sequence by *in vivo* homologous recombination, approximately 40 ng of each plasmid linearized by digestion with BssHII and 1 ng of the pCLS0542 (SEQ ID NO: 156) plasmid DNA linearized by digestion with NcoI and EagI were used to transform, respectively, the yeast *S. cerevisiae* strain FYC2-6A (*MAT*α, *trp1*Δ63, *leu2*Δ1, *his3*Δ200) using a high efficiency LiAc transformation protocol (Arnould et al. 2007).

All the yeast target reporter plasmids containing the TALEN DNA target sequences were constructed as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

The TALE-AvrBs3::TevI constructs were tested in a yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006) on pseudo palindromic targets in order to compare activity with a standard TALE-AvrBs3::FokI TALEN (pCLS8590, SEQ ID NO: 244), which requires two binding sites for activity. AvrBs3 targets contain two identical recognition sequences juxtaposed with the 3' ends proximal and separated by "spacer" DNA ranging from 5 to

40 bps (SEQ ID NO: 157 to 192, Table 6). TALE-AvrBs3::TevI activity levels on their respective targets in yeast cells are shown on figure 9. Data summarized in figure 9 show that TALE-AvrBs3::TevI is active against several targets in Yeast.

Activity of TALE::TevI in mammalian cells

DNA encoding the TALE-AvrBs3::TevI construct from either pCLS9012 (SEQ ID NO: 218) or pCLS9013 (SEQ ID NO: 153) was subcloned into the pCLS1853 (SEQ ID NO: 193) mammalian expression plasmid using AscI and XhoI restriction enzymes for the receiving plasmid and BssHII and XhoI restriction enzymes for the TALE-AvrBs3::TevI insert, leading to the mammalian expression plasmids pCLS8993 and pCLS8994 (SEQ ID NO: 194 and 195), respectively.

All mammalian target reporter plasmids containing the TALEN DNA target sequences were constructed using the standard Gateway protocol (INVITROGEN) into a CHO reporter vector (Arnould, Chames et al. 2006, Grizot, Epinat et al. 2010). The TALE-AvrBs3::TevI constructs were tested in an extrachromosomal assay in mammalian cells (CHO K1) on pseudo palindromic targets in order to compare activity with a standard TALE-AvrBs3::FokI TALEN, which requires two binding sites for activity. AvrBs3 targets contain two identical recognition sequences juxtaposed with the 3' ends proximal and separated by "spacer" DNA ranging from 5 to 40 bps (SEQ ID NO: 157 to 192, Table 6).

For this assay, CHO K1 cells were transfected in a 96-well plate format with 75 ng of target vector and an increasing quantity of each variant DNA from 0.7 to 25 ng , in the presence of PolyFect reagent (1µL per well). The total amount of transfected DNA was completed to 125 ng (target DNA, variant DNA, carrier DNA) using an empty vector. Seventy-two hours after transfection, culture medium was removed and 150µl of lysis/revelation buffer for β-galactosidase liquid assay was added. After incubation at 37°C, optical density was measured at 420 nm. The entire process is performed on an automated Velocity11 BioCel platform (Grizot, Epinat et al. 2009).

Activity levels in mammalian cells for the TALE-AvrBs3::TevI constructs (12.5 ng DNA transfected) on the Avr15 target (SEQ ID NO: 167) are shown in figure 10. TALE-AvrBs3::TevI appears to be efficient to cleave the target sequence.

TALE::TevI nickase activity

The results described in examples above illustrate two TALE::TevI fusions, each containing one TALE-based DNA binding domain and one I-TevI-based catalytic domain, working to generate detectable activity. The assays used measure tandem repeat recombination by single-strand annealing, a

process that is triggered essentially by a DSB (Sugawara and Haber 1992; Paques and Duchateau 2007). TALE::TevI fusions can have a nickase activity insufficient to alone trigger a signal in the cell-based assay. However, two TALE::TevI proteins binding on two nearby sites can sometimes generate two independent nicks, that when proximal and on different DNA strands can create a DSB. In this case, each TALE::TevI is a cTALEN able to generate a nick.

Different experiments are set up to measure TALE::TevI nickase activity:

Super-coiled circular plasmid nicking and / or linearization assay

The sequences encoding the TALE-AvrBs3::TevI constructs cT11Avr_TevD01 and cT11Avr_TevD02 are cloned into a T7-based expression vector using NcoI/EagI restriction sites to yield plasmids pCLS9021 (SEQ ID NO: 201) and pCLS9022 (SEQ ID NO: 202), respectively. This cloning step results in TALE-AvrBs3::TevI proteins having an additional hexa-His tag for purification. Plasmids pCLS9021 and pCLS9022 are then used to produce active proteins by one of two methods:

1. Plasmids are used in a standard *in vitro* transcription/translation system; lysates from the translation are used directly without further purification.

2. Plasmids are used to transform E.coli BL21(DE3) cells for expression using standard protocols, namely: growth to log phase, induction with IPTG, harvest, cell lysis and purification via affinity methods for His-tagged proteins.

Active proteins are assayed against DNA targets having either none, one or two AvrBs3 recognition site sequences. When more than one site is present, identical recognition sequences are juxtaposed with the 3' ends proximal and separated by "spacer" DNA ranging from 5 to 40 bps.

A super-coiled circular plasmid nicking and/or linearization assay is performed. Plasmids harboring the DNA targets described above are prepared by standard methods and column purified to yield super-coiled plasmid of >98% purity. Increasing amounts of TALE-AvrBs3::TevI proteins (prepared as described above) are incubated with each plasmid under conditions to promote DNA cleavage for 1h at 37°C. Reaction products are separated on agarose gels and visualized by EtBr staining.

Linear DNA nicking and/or cleavage assay

A linear DNA nicking and/or cleavage assay is also performed. PCR products containing the target sequences described above are prepared by standard methods and column purified to yield linear substrate of >98% purity. Increasing amounts of TALE-AvrBs3::TevI proteins (prepared as described

above) are then incubated with each PCR substrate under conditions to promote DNA cleavage for 1h at 37°C. Reaction products are separated on a denaturing acrylamide gel and the single-strand DNA visualized.

Engineering of the TALE::TevI

Variants differing by truncations of the C-terminal domain of the AvrBs3-derived TALEN (SEQ ID NO: 196) are chosen as starting scaffolds. A subset of these variants includes truncation after positions E886 (C0), P897 (C11), G914 (C28), L926 (C40), D950 (C64), R1000 (C115), D1059 (C172) (the protein domains of truncated C-terminal domains C11 to C172 are respectively given in SEQ ID NO: 204 to 209) and P1117 [also referred as Cter wt or WT Cter (SEQ ID NO: 210) lacking the activation domain of the C-terminal domain of natural AvrBs3 (SEQ ID NO: 220)]. The plasmids coding for the variant scaffolds containing the AvrBs3-derived N-terminal domain, the AvrBs3-derived set of repeat domains and the truncated AvrBs3-derived C-terminal domain [pCLS7821, pCLS7803, pCLS7807, pCLS7809, pCLS7811, pCLS7813, pCLS7817 (SEQ ID NO: 211 to 217) which are based on the pCLS7184 (SEQ ID NO: 196)] allow cloning of any catalytic domain in fusion to the C-terminal domain, using the restriction sites BamHI and EagI.

Variants of the catalytic domain of I-TevI (SEQ ID NO: 20) are designed from the N-terminal region of I-TevI. A subset of these variants includes truncations of the catalytic domain, as the deletion-intolerant region of its linker, the deletion-tolerant region of its linker and its zinc finger (SEQ ID NO: 197 to 200) named in Liu et al, 2008 (Liu, Dansereau et al. 2008).

The DNA corresponding to these variants of I-TevI is amplified by the PCR to introduce, at the DNA level, a BamHI (at the 5' of the coding strand) and a EagI (at the 3' of the coding strand) restriction site and, at the protein level, a linker (for example -SGGSGS- stretch, SEQ ID NO: 219) between the C terminal domain of the TALE and the variant of the catalytic domain of I-TevI. The final TALE::TevI constructs are generated by insertion of the variant of I-TevI catalytic domains into the scaffold variants using BamHI and EagI and standard molecular biology procedures.

All the yeast target reporter plasmids containing the TALEN DNA target sequences were constructed as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

The TALE-AvrBs3::TevI constructs were tested in a yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006) on pseudo palindromic targets in

order to compare activity with a standard TALE-AvrBs3::FokI TALEN (pCLS8590, SEQ ID NO: 244), which requires two binding sites for activity. AvrBs3 targets contain two identical recognition sequences juxtaposed with the 3' ends proximal and separated by "spacer" DNA ranging from 5 to 40 bps (SEQ ID NO: 157 to 192, Table 6).

Example 3b: TevI::TALE compact TALEN

The sT2 (SEQ ID NO: 135) core TALE scaffold described in example 3a was selected to generate pCLS7865-cTAL11_NFS1 (pCLS9008, SEQ ID NO: 234), where NFS1 designates the amino acid sequence -GSSG- (with underlying restriction sites BamHI and Kpn2I in the coding DNA to facilitate cloning). Four variants of the I-TevI (SEQ ID NO: 20) catalytic domain were amplified by the PCR on templates TevCreD01 [SEQ ID NO: 109 protein in plasmid pCLS6614 (SEQ ID NO: 146)] using the primer pairs CMP_G001 (SEQ ID NO: 239) and CMP_G067 (SEQ ID NO: 263) or CMP_G152 (SEQ ID NO: 264), TevCreD02 [SEQ ID NO: 110 protein in plasmid pCLS6615 (SEQ ID NO: 203)] using the primer pair CMP_G001 (SEQ ID NO: 239) and CMP_G068 (SEQ ID NO: 240) or TevCreD05 [SEQ ID NO: 113 protein in plasmid pCLS6618 (SEQ ID NO: 258)] using the primer pair CMP_G001 (SEQ ID NO: 239) and CMP_G114 (SEQ ID NO: 265) and subcloned into the pCLS9008 backbone by restriction and ligation using NcoI and Kpn2I restriction sites, yielding pCLS7865-TevW01_cT11 (pCLS15777, SEQ ID NO: 266, encoding the protein of SEQ ID NO: 426), pCLS7865-TevD01_cT11 (pCLS15778, SEQ ID NO: 267, encoding the protein of SEQ ID NO: 427), pCLS7865-TevD02_cT11 (pCLS12730, SEQ ID NO: 235, encoding the protein of SEQ ID NO: 428) and pCLS7865-TevD05_cT11 (pCLS15779, SEQ ID NO: 268 , encoding the protein of SEQ ID NO: 429), respectively. Whereas the TevW01_cT11-based fusion contains the dipeptide -SG- linking the TALE-derived DNA binding domain and I-TevI-derived catalytic domain, all others constructs incorporate a longer pentapeptide -QGPSG- to link the domains.

Activity of TevI::TALE in yeast

The DNA sequence coding for the RVDs to target the AvrBs3 site (SEQ ID NO: 152) was subcloned into plasmids pCLS15777 (SEQ ID NO: 266), pCLS15778 (SEQ ID NO: 267) and pCLS12730 (SEQ ID NO: 235) using Type IIS restriction enzymes BsmBI for the receiving plasmid and BbvI and SfaNI for the inserted RVD sequence to create the subsequent TevI::TALE-AvrBs3 constructs TevW01_cT11Avr (pCLS15780, SEQ ID NO: 269, encoding the protein of SEQ ID NO: 430), TevD01_cT11Avr (pCLS15781, SEQ ID NO: 270, encoding the protein of SEQ ID NO: 431) and TevD02_cT11Avr (pCLS12731, SEQ ID NO: 236, encoding the protein of SEQ ID NO: 432), respectively. A similar cloning technique was used to introduce the RVDs to target the RagT2-R site (SEQ ID NO: 271) into plasmid pCLS15779 (SEQ ID

NO: 268) to create the subsequent construct TevD05_cT11RagT2-R (pCLS15782, SEQ ID NO: 272). All TevI::TALE constructs were sequenced and the inserts transferred to additional vectors as needed (see below).

The final TevI::TALE-based yeast expression plasmids, pCLS11979 (SEQ ID NO: 273), pCLS8521 (SEQ ID NO: 274), pCLS8522 (SEQ ID NO: 237) and pCLS12100 (SEQ ID NO: 275), were prepared by yeast *in vivo* cloning using plasmid pCLS15780 (SEQ ID NO: 269), pCLS15781 (SEQ ID NO: 270), pCLS12731 (SEQ ID NO: 236) and pCLS15782 (SEQ ID NO: 272), respectively. To generate an intact coding sequence by *in vivo* homologous recombination, approximately 40 ng of each plasmid linearized by digestion with BssHII and 1 ng of the pCLS0542 (SEQ ID NO: 156) plasmid DNA linearized by digestion with NcoI and EagI were used to transform, respectively, the yeast *S. cerevisiae* strain FYC2-6A (*MATα*, *trp1Δ63*, *leu2Δ1*, *his3Δ200*) using a high efficiency LiAc transformation protocol (Arnould et al. 2007).

All the yeast target reporter plasmids containing the TALEN DNA target sequences were constructed as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

The TevI::TALE-AvrBs3 and TevI::TALE-RagT2-R constructs were tested in a yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006) on pseudo palindromic targets in order to compare activity with a standard TALE-AvrBs3::FokI TALEN (pCLS8590, SEQ ID NO: 244), which requires two binding sites for activity. AvrBs3 targets contain two identical recognition sequences juxtaposed with the 3' ends proximal and separated by "spacer" DNA ranging from 5 to 40 bps (SEQ ID NO: 157 to 192, Table 6). In addition, constructs were tested on a target having only a single AvrBs3 or RagT2-R recognition site (SEQ ID NO: 238, Table 8). The TevI::TALE-AvrBs3 activity level in yeast was comparable to that of TALE-AvrBs3::TevI (pCLS8524, SEQ ID NO: 155) on suitable targets. Significant activity is illustrated in table 8 for a sample single-site target, according to the cTALEN of the present invention.

	TALEN Construct		
Target DNA	TevI::TALE-AvrBs3	TevI::TALE-RagT2-R	TALE-AvrBs3::FokI
Avr25 (dual-site) (SEQ ID NO: 177)	++++	n.d.	++++
Avr25RAGT2R (single-site)(SEQ ID NO: 238)	++	++	n.d.

Table 8: Activity of TevI::TALE-AvrBs3 and and TevI::TALE-RagT2-R on dual- and single-site DNA targets. Relative activity is scaled as: n.d., no activity detectable; +, <25% activity; ++, 25% to <50% activity; +++, 50% to <75% activity; ++++, 75% to 100% activity.

Activity of TevI::TALE in plants

The DNA sequence coding for the RVDs to target the NptIIT5-L and NptIIT6-L sites (SEQ ID NO: 276 to 279) were subcloned into plasmid pCLS12730 (SEQ ID NO: 235) using Type IIS restriction enzymes BsmBI for the receiving plasmid and BbvI and SfaNI for the inserted RVD sequences to create the subsequent TevI::TALE constructs TevD02_cT11NptIIT5-L (pCLS15783, SEQ ID NO: 280) and TevD02_cT11NptIIT6-L (pCLS15784, SEQ ID NO: 281), respectively. The constructs were sequenced and the TevI::TALE inserts transferred by standard cloning techniques to plasmid pCLS14529 (SEQ ID NO: 282) to generate the final TevI::TALE-NptIIT5-L and TevI::TALE-NptIIT6-L expression plasmids, pCLS14579 (SEQ ID NO: 283) and pCLS14581 (SEQ ID NO: 284), respectively. Plasmid pCLS14529 allows for cloning gene of interest sequences downstream of a promoter that confers high levels of constitutive expression in plant cells.

To test activity in plant cells, a YFP-based single-strand annealing (SSA) assay was employed. The YFP reporter gene has a short duplication of coding sequence that is interrupted by either an NptIIT5 or NptIIT6 TALEN target site. Cleavage at the target site stimulates recombination between the repeats, resulting in reconstitution of a functional YFP gene. To quantify cleavage, the reporter is introduced along with a construct encoding a FokI-based TALEN or compact TALEN into tobacco protoplasts by PEG-mediated transformation (as known or derived from the state of the art). Uniform transformation efficiencies were obtained by using the same amount of plasmid in each transformation – i.e. 15 μg each of plasmids encoding YFP and either the TALEN or cTALEN. After 24 hours, the protoplasts were subjected to flow cytometry to quantify the number of YFP positive cells. The TevI::TALE activity levels, using cTALENs according to the present invention, in plants were comparable to those of a FokI-based TALEN control constructs on the targets tested (Table 9).

	TALEN Construct			
Target DNA	TevI::TALE-NptIIT5-L	NptII5.1 control	TevI::TALE-NptIIT6-L	NptII6.1 control
NptII5.1	+++	+	n.a.	n.a.
NptII6.1	n.a.	n.a.	+	+

Table 9: Activity of TevI::TALE-NptIIT5-L and TevI::TALE-NptIIT6-L on appropriate DNA targets. Relative activity is scaled to the control constructs as: n.a., not applicable; +, 100% activity of control (2% YFP positive cells).

Example 3c: TALE::NucA compact TALEN

NucA (SEQ ID NO: 26), a nonspecific endonuclease from *Anabaena* sp., was fused to a TALE-derived scaffold (composed of a N-terminal domain, a central core composed of RVDs and a C-terminal domain) to create a new class of cTALEN (TALE::NucA). To distinguish the orientation (N-terminal vs. C-terminal) of the catalytic domain (CD) fusions, construct names are written as either CD::TALE-RVD (catalytic domain is fused N-terminal to the TALE domain) or TALE-RVD::CD (catalytic domain is fused C-terminal to the TALE domain), where "-RVD" optionally designates the sequence recognized by the TALE domain and "CD" is the catalytic domain type. Herein, we describe novel TALE::NucA constructions that target for example the AvrBs3 sequence, and are thus named TALE-AvrBs3::NucA. Notably, the wild-type NucA endonuclease can be inhibited by complex formation with the NuiA protein (SEQ ID NO: 229). In a compact TALEN context, the NuiA protein can function as an auxiliary domain to modulate the nuclease activity of TALE::NucA constructs.

Activity of TALE::NucA in yeast

A core TALE scaffold, sT2 (SEQ ID NO: 135), was selected onto which (a) different sets of RVD domains could be inserted to change DNA binding specificity, and; (b) a selection of NucA-derived catalytic domains could be attached, N- or C-terminal, to effect DNA cleavage (or nicking). As previously mentioned, the sT2 truncated scaffold was generated by the PCR from a full-length core TALEN scaffold template (pCLS7183, SEQ ID NO: 141) using primers CMP_G061 (SEQ ID NO: 142) and CMP_G065 (SEQ ID NO: 143) and was cloned into vector pCLS7865 (SEQ ID NO: 144) to generate pCLS7865-cTAL11_CFS1 (pCLS9009, SEQ ID NO: 145), where CFS1 designates the amino acid sequence -GSSG- (with underlying restriction sites BamHI and Kpn2I in the coding DNA to facilitate cloning). The NucA (SEQ ID NO: 26) catalytic domain, corresponding to amino acid residues 25 to 274, was subcloned into the pCLS9009 backbone (SEQ ID NO: 145) by restriction and ligation using BamHI and EagI restriction sites, yielding pCLS7865-cT11_NucA (pCLS9937, SEQ ID NO: 221, encoding the protein of SEQ ID NO: 433). The fusion contains the dipeptide -GS- linking the TALE-derived DNA binding domain and NucA-derived catalytic domain. The cloning step also brings at the amino acid level an AAD sequence at the Cter of the NucA catalytic domain.

The DNA sequence coding for the RVDs to target the AvrBs3 site (SEQ ID NO: 152) was subcloned into plasmid pCLS9937 (SEQ ID NO: 221) using Type IIS restriction enzymes BsmBI for the receiving plasmid and BbvI and SfaNI for the inserted RVD sequence to create the subsequent TALE-

AvrBs3::NucA construct cT11Avr_NucA (pCLS9938, SEQ ID NO: 222, encoding the protein of SEQ ID NO: 434). The TALE-AvrBs3::NucA construct was sequenced and the insert transferred to additional vectors as needed (see below).

The final TALE-AvrBs3::NucA yeast expression plasmid, pCLS9924 (SEQ ID NO: 223), was prepared by yeast *in vivo* cloning using plasmid pCLS9938 (SEQ ID NO: 222). To generate an intact coding sequence by *in vivo* homologous recombination, approximately 40 ng of plasmid (pCLS9938) linearized by digestion with BssHII and 1 ng of the pCLS0542 (SEQ ID NO: 156) plasmid DNA linearized by digestion with NcoI and EagI were used to transform the yeast *S. cerevisiae* strain FYC2-6A (*MATα*, *trp1*Δ63, *leu2*Δ1, *his3*Δ200) using a high efficiency LiAc transformation protocol (Arnould et al. 2007).

All the yeast target reporter plasmids containing the TALEN DNA target sequences were constructed as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

The TALE-AvrBs3::NucA construct was tested in a yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006) on pseudo palindromic targets in order to compare activity with a standard TALE-AvrBs3::FokI TALEN (pCLS8590, SEQ ID NO: 244), which requires two binding sites for activity. AvrBs3 targets contain two identical recognition sequences juxtaposed with the 3' ends proximal and separated by "spacer" DNA ranging from 5 to 40 bps (SEQ ID NO: 157 to 192, Table 7). In addition, constructs were tested on a target having only a single AvrBs3 recognition site (SEQ ID NO: 224; Table 7).

Engineering of the TALE::NucA

Variants differing by truncations of the C-terminal domain of the AvrBs3-derived TALEN (SEQ ID NO: 196) are chosen as starting scaffolds. A subset of these variants includes truncation after positions E886 (C0), P897 (C11), G914 (C28), L926 (C40), D950 (C64), R1000 (C115), D1059 (C172) (the protein domains of truncated C-terminal domains C11 to C172 are respectively given in SEQ ID NO: 204 to 209) and P1117 [also referred as Cter wt or WT Cter (SEQ ID NO: 210) lacking the activation domain of the C-terminal domain of natural AvrBs3 (SEQ ID NO: 220)]. The plasmids coding for the variant scaffolds containing the AvrBs3-derived N-terminal domain, the AvrBs3-derived set of repeat domains and the truncated AvrBs3-derived C-terminal domain [pCLS7821, pCLS7803, pCLS7807, pCLS7809, pCLS7811, pCLS7813, pCLS7817 (SEQ ID NO: 211 to 217) which are based on the

pCLS7184 (SEQ ID NO: 196)] allow cloning of any catalytic domain in fusion to the C-terminal domain, using the restriction sites BamHI and EagI.

The DNA corresponding to amino acid residues 25 to 274 of NucA is amplified by the PCR to introduce, at the DNA level, a BamHI (at the 5' of the coding strand) and a EagI (at the 3' of the coding strand) restriction site and, at the protein level, a linker (for example -SGGSGS- stretch, SEQ ID NO: 219) between the C terminal domain of the TALE and the NucA catalytic domain. The final TALE::NucA constructs are generated by insertion of the NucA catalytic domain into the scaffold variants using BamHI and EagI and standard molecular biology procedures. For example, scaffold variants truncated after positions P897 (C11), G914 (C28) and D950 (C64), respectively encoded by pCLS7803, pCLS7807, pCLS7811, (SEQ ID NO: 212, 213 and 215), were fused to the NucA catalytic domain (SEQ ID NO: 26), leading to pCLS9596, pCLS9597, and pCLS9599 (SEQ ID NO: 225 to 227). The cloning step also brings at the amino acid level an AAD sequence at the Cter of the NucA catalytic domain.

All the yeast target reporter plasmids containing the TALEN DNA target sequences were constructed as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

The TALE-AvrBs3::NucA constructs were tested in a yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006) on pseudo palindromic targets in order to compare activity with a standard TALE-AvrBs3::FokI TALEN (pCLS8590, SEQ ID NO: 244), which requires two binding sites for activity. AvrBs3 targets contain two identical recognition sequences juxtaposed with the 3' ends proximal and separated by "spacer" DNA ranging from 5 to 40 bps (SEQ ID NO: 157 to 192, Table 7). In addition, TALE-AvrBs3::NucA constructs were tested on a target having only a single AvrBs3 recognition site (SEQ ID NO: 224). Data summarized in figure 11 show that TALE-AvrBs3::NucA constructs are active on all targets having at least one AvrBs3 recognition site, according to the cTALEN of the present invention.

Example 3d: TALE::ColE7 compact TALEN

The catalytic domain of ColE7 (SEQ ID NO: 140), a nonspecific endonuclease from *E.coli*, was fused to a TALE-derived scaffold (composed of a N-terminal domain, a central core composed of RVDs and a C-terminal domain) to create a new class of cTALEN (TALE::ColE7). To distinguish the orientation (N-

terminal vs. C-terminal) of the catalytic domain (CD) fusions, construct names are written as either CD::TALE-RVD (catalytic domain is fused N-terminal to the TALE domain) or TALE-RVD::CD (catalytic domain is fused C-terminal to the TALE domain), where "-RVD" optionally designates the sequence recognized by the TALE domain and "CD" is the catalytic domain type. Herein, we describe novel TALE::ColE7 constructions that target for example the AvrBs3 sequence, and are thus named TALE-AvrBs3::ColE7. Notably, the wild-type ColE7 endonuclease can be inhibited by complex formation with the Im7 immunity protein (SEQ ID NO: 230). In a compact TALEN context, the Im7 protein can function as an auxiliary domain to modulate the nuclease activity of TALE::ColE7 constructs.

Activity of TALE::ColE7 in yeast

A core TALE scaffold, sT2 (SEQ ID NO: 135), was selected onto which (a) different sets of RVD domains could be inserted to change DNA binding specificity, and; (b) a selection of ColE7-derived catalytic domains could be attached, N- or C-terminal, to effect DNA cleavage (or nicking). As previously mentioned, the sT2 truncated scaffold was generated by the PCR from a full-length core TALEN scaffold template (pCLS7183, SEQ ID NO: 141) using primers CMP_G061 (SEQ ID NO: 142) and CMP_G065 (SEQ ID NO: 143) and was cloned into vector pCLS7865 (SEQ ID NO: 144) to generate pCLS7865-cTAL11_CFS1 (pCLS9009, SEQ ID NO: 145), where CFS1 designates the amino acid sequence -GSSG- (with underlying restriction sites BamHI and Kpn2I in the coding DNA to facilitate cloning). The ColE7 (SEQ ID NO: 140) catalytic domain was subcloned into the pCLS9009 backbone by restriction and ligation using Kpn2I and EagI restriction sites, yielding pCLS7865-cT11_ColE7 (pCLS9939, SEQ ID NO: 231, encoding the protein of SEQ ID NO: 435). The fusion contains the dipeptide -GSSG- linking the TALE-derived DNA binding domain and ColE7-derived catalytic domain.

The DNA sequence coding for the RVDs to target the AvrBs3 site (SEQ ID NO: 152) was subcloned into plasmid pCLS9939 (SEQ ID NO: 231) using Type IIS restriction enzymes BsmBI for the receiving plasmid and BbvI and SfaNI for the inserted RVD sequence to create the subsequent TALE-AvrBs3::ColE7 construct cT11Avr_ColE7 (pCLS9940, SEQ ID NO: 232, encoding the protein of SEQ ID NO: 436). The TALE-AvrBs3::ColE7 construct was sequenced and the insert transferred to additional vectors as needed (see below).

The final TALE-AvrBs3::ColE7 yeast expression plasmid, pCLS8589 (SEQ ID NO: 233), was prepared by yeast *in vivo* cloning using plasmid pCLS9940 (SEQ ID NO: 232). To generate an intact coding sequence by *in vivo* homologous recombination, approximately 40 ng of plasmid (pCLS9940) linearized by digestion with BssHII and 1 ng of the pCLS0542 (SEQ ID NO: 156) plasmid DNA linearized by digestion with NcoI and EagI were used to transform the yeast *S. cerevisiae* strain FYC2-

6A (*MATα*, *trp1*Δ63, *leu2*Δ1, *his3*Δ200) using a high efficiency LiAc transformation protocol (Arnould et al. 2007).

All the yeast target reporter plasmids containing the TALEN DNA target sequences were constructed as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

The TALE-AvrBs3::ColE7 construct was tested in a yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006) on pseudo palindromic targets in order to compare activity with a standard TALE-AvrBs3::FokI TALEN (pCLS8590, SEQ ID NO: 244), which requires two binding sites for activity. AvrBs3 targets contain two identical recognition sequences juxtaposed with the 3' ends proximal and separated by "spacer" DNA ranging from 5 to 40 bps (SEQ ID NO: 157 to 192, Table 7). In addition, constructs were tested on a target having only a single AvrBs3 recognition site (SEQ ID NO: 224, Table 7). TALE-AvrBs3::ColE7 activity levels on the respective targets in yeast cells are shown in figure 12.

Activity of TALE::ColE7 in plants

The DNA sequence coding for the RVDs to target the NptIIT5-L and NptIIT6-L sites (SEQ ID NO: 276 to 279) were subcloned into plasmid pCLS15785 (SEQ ID NO: 285, a C-terminally modified ColE7 K497A mutant of plasmid pCLS9939, SEQ ID NO: 231) using Type IIS restriction enzymes BsmBI for the receiving plasmid and BbvI and SfaNI for the inserted RVD sequences to create the subsequent TALE::ColE7_A497 constructs cT11NptIIT5-L_ColE7_A497 (pCLS15786, SEQ ID NO: 286) and cT11NptIIT6-L_ColE7_A497 (pCLS15787, SEQ ID NO: 287), respectively. The constructs were sequenced and the TALE::ColE7_A497 inserts transferred by standard cloning techniques to plasmid pCLS14529 (SEQ ID NO: 282) to generate the final TALE-NptIIT5-L::ColE7_A497 and TALE-NptIIT6-L::ColE7_A497 expression plasmids, pCLS14584 (SEQ ID NO: 288, encoding the protein of SEQ ID NO: 437) and pCLS14587 (SEQ ID NO: 289, encoding the protein of SEQ ID NO: 438), respectively. Plasmid pCLS14529 allows for cloning gene of interest sequences downstream of a promoter that confers high levels of constitutive expression in plant cells.

To test activity in plant cells, a YFP-based single-strand annealing (SSA) assay was employed. The YFP reporter gene has a short duplication of coding sequence that is interrupted by either an NptIIT5 or NptIIT6 TALEN target site. Cleavage at the target site stimulates recombination between the repeats, resulting in reconstitution of a functional YFP gene. To quantify cleavage, the reporter

is introduced along with a construct encoding a FokI-based TALEN or compact TALEN into tobacco protoplasts by PEG-mediated transformation. Uniform transformation efficiencies were obtained by using the same amount of plasmid in each transformation – i.e. 15 μg each of plasmids encoding YFP and either the TALEN or cTALEN. After 24 hours, the protoplasts were subjected to flow cytometry to quantify the number of YFP positive cells. The TALE:: ColE7_A497 activity levels, using cTALENs according to the present invention, in plants were comparable to those of a FokI-based TALEN control constructs on the targets tested (Table 10).

	TALEN Construct			
Target DNA	TALE-NptIIT5-L::ColE7_A497	NptII5.1 control	TALE-NptIIT6-L::ColE7_A497	NptII6.1 control
NptII5.1	+	+	n.a.	n.a.
NptII6.1	n.a.	n.a.	+	+

Table 10: Activity of TALE-NptIIT5-L::ColE7_A497 and TALE-NptIIT6-L::ColE7_A497 on appropriate DNA targets. Relative activity is scaled to the control constructs as: n.a., not applicable; +, 100% activity of control (8% YFP positive cells).

Engineering of the TALE::ColE7

Variants differing by truncations of the C-terminal domain of the AvrBs3-derived TALEN (SEQ ID NO: 196) are chosen as starting scaffolds. A subset of these variants includes truncation after positions E886 (C0), P897 (C11), G914 (C28), L926 (C40), D950 (C64), R1000 (C115), D1059 (C172) (the protein domains of truncated C-terminal domains C11 to C172 are respectively given in SEQ ID NO: 204 to 209) and P1117 [also referred as Cter wt or WT Cter (SEQ ID NO: 210) lacking the activation domain of the C-terminal domain of natural AvrBs3 (SEQ ID NO: 220)]. The plasmids coding for the variant scaffolds containing the AvrBs3-derived N-terminal domain, the AvrBs3-derived set of repeat domains and the truncated AvrBs3-derived C-terminal domain [pCLS7821, pCLS7803, pCLS7807, pCLS7809, pCLS7811, pCLS7813, pCLS7817 (SEQ ID NO: 211 to 217) which are based on the pCLS7184 (SEQ ID NO: 196)] allow cloning of any catalytic domain in fusion to the C-terminal domain, using the restriction sites BamHI and EagI.

The DNA corresponding to the catalytic domain of ColE7 is amplified by the PCR to introduce, at the DNA level, a BamHI (at the 5' of the coding strand) and a EagI (at the 3' of the coding strand) restriction site and, at the protein level, a linker (for example -SGGSGS- stretch, SEQ ID NO: 219) between the C terminal domain of the TALE and the ColE7 catalytic domain. Additionally, variants of the ColE7 endonuclease domain that modulate catalytic activity can be generated having changes (individually or combined) at the following positions: K446, R447, D493,

R496, K497, H545, N560 and H573 [positions refer to the amino acid sequence of the entire ColE7 protein (SEQ ID NO: 11)]. The final TALE::ColE7 constructs are generated by insertion of the ColE7 catalytic domain into the scaffold variants using BamHI and EagI and standard molecular biology procedures.

All the yeast target reporter plasmids containing the TALEN DNA target sequences were constructed as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

The TALE-AvrBs3::ColE7 constructs are tested in a yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006) on pseudo palindromic targets in order to compare activity with a standard TALE-AvrBs3::FokI TALEN (pCLS8590, SEQ ID NO: 244), which requires two binding sites for activity. AvrBs3 targets contain two identical recognition sequences juxtaposed with the 3' ends proximal and separated by "spacer" DNA ranging from 5 to 40 bps (SEQ ID NO: 157 to 192, Table 7). In addition, constructs were tested on a target having only a single AvrBs3 recognition site (SEQ ID NO: 224, Table 7).

Example 3e: TALE::CreI compact TALEN

The wild-type I-CreI meganuclease (SEQ ID NO: 106) was chosen as a protein template to derive a sequence-specific catalytic domain that when fused to a TALE-derived scaffold (composed of a N-terminal domain, a central core composed of RVDs and a C-terminal domain) would generate a new class of cTALEN (TALE::CreI). To distinguish the orientation (N-terminal vs. C-terminal) of the catalytic domain (CD) fusions, construct names are written as either CD::TALE-RVD (catalytic domain is fused N-terminal to the TALE domain) or TALE-RVD::CD (catalytic domain is fused C-terminal to the TALE domain), where "-RVD" optionally designates the sequence recognized by the TALE domain and "CD" is the catalytic domain type. Herein, we describe novel TALE::CreI-based constructions that target for example the T cell receptor B gene (TCRB gene, SEQ ID NO: 290, Figure 13) sequence, both via the TALE DNA binding domain and the re-engineered I-CreI domain. Notably, specificity of the TALE::CreI compact TALEN is driven by both the TALE DNA binding domain as well as the I-CreI-derived catalytic domain. In a compact TALEN context, such proteins can provide, within a reasonably-sized monomeric protein, the requisite high specificity demanded by therapeutic applications.

Activity of TALE::CreI in yeast

A core TALE scaffold, sT2 (SEQ ID NO: 135), was selected onto which (a) different sets of RVD domains could be inserted to change DNA binding specificity, and; (b) a selection of I-CreI-derived catalytic domains could be attached, N- or C-terminal, to effect DNA cleavage (or nicking). As previously mentioned, the sT2 truncated scaffold was generated by the PCR from a full-length core TALEN scaffold template (pCLS7183, SEQ ID NO: 141) using primers CMP_G061 (SEQ ID NO: 142) and CMP_G065 (SEQ ID NO: 143) and was cloned into vector pCLS7865 (SEQ ID NO: 144) to generate pCLS7865-cTAL11_CFS1 (pCLS9009, SEQ ID NO: 145), where CFS1 designates the amino acid sequence -GSSG- (with underlying restriction sites BamHI and Kpn2I in the coding DNA to facilitate cloning). A re-engineered I-CreI catalytic domain, designed to target a sequence in the T cell receptor B gene (TCRB gene, SEQ ID NO: 290, Figure 13), was subcloned in two steps. First, the I-CreI_NFS1 (SEQ ID NO: 122) scaffold, where NFS1 (SEQ ID NO: 98) comprises a linker of 20 amino acids -GSDITKSKISEKMKGQGPSG- (with underlying restriction sites BamHI and Kpn21 in the coding DNA to facilitate cloning), was fused to the pCLS7865-cTAL11_CFS1 scaffold (using BamHI and EagI restriction sites) to insert the NFS1 linker in-frame to the coding sequence. The I-CreI meganuclease was subsequently replaced by the engineered TCRB02-A meganuclease (pCLS6857, SEQ ID NO: 291) construct using Kpn2I and XhoI restriction sites, yielding pCLS7865-cT11_scTB2aD01 (pCLS15788, SEQ ID NO: 292, encoding the protein of SEQ ID NO: 439). Two point-mutant variants of the TCRB02-A meganuclease, TCRB02-A_148C (pCLS12083, SEQ ID NO: 293, encoding the protein of SEQ ID NO: 442) and TCRB02-A_333C (pCLS12195, SEQ ID NO: 294, encoding the protein of SEQ ID NO: 443), were also subcloned as catalytic domains fused to a TALE binding core, yielding constructs pCLS7865-cT11_scTB2aD01_148C (pCLS15789, SEQ ID NO: 295, encoding the protein of SEQ ID NO: 440) and pCLS7865-cT11_scTB2aD01_333C (pCLS15790, SEQ ID NO: 296, encoding the protein of SEQ ID NO: 441).

Three DNA sequences coding for RVDs that target the TCRB gene were designed at different distances from the meganuclease site, leading to RVDs TCRB02A1 (SEQ ID NO: 297), TCRB02A2 (SEQ ID NO: 298) and TCRB02A3 (SEQ ID NO: 299) that target sequences located 7 bp, 12 bp and 16 bp, respectively, upstream of the meganuclease TCRB site (Figure 13). DNA sequences for each RVD were independently subcloned into plasmid pCLS15788 (SEQ ID NO: 292) using Type IIS restriction enzymes BsmBI for the receiving plasmid and BbvI and SfaNI for the inserted RVD sequence to create the subsequent TALE::scTB2aD01 constructs cT11TB2A1_scTB2aD01 (pCLS15791, SEQ ID NO: 300), cT11TB2A2_scTB2aD01 (pCLS15792, SEQ ID NO: 301) and cT11TB2A3_scTB2aD01 (pCLS15793, SEQ ID NO: 302). Additionally, the TCRB02A2 (SEQ ID NO: 298) RVDs were similarly cloned into pCLS15789 (SEQ ID NO: 295) to create cT11TB2A2_scTB2aD01_148C (pCLS15794, SEQ ID NO: 303).

All constructs were sequenced and the various inserts transferred to additional vectors as needed (see below).

The final TALE::scTB2aD01 yeast expression plasmids, pCLS13449 (SEQ ID NO: 304, encoding the protein of SEQ ID NO: 444), pCLS13450 (SEQ ID NO: 305, encoding the protein of SEQ ID NO: 445), pCLS13451 (SEQ ID NO: 306, encoding the protein of SEQ ID NO: 446) and pCLS15148 (SEQ ID NO: 307, encoding the protein of SEQ ID NO: 455), were prepared by yeast *in vivo* cloning using plasmids pCLS15791 (SEQ ID NO: 300), pCLS15792 (SEQ ID NO: 301), pCLS15793 (SEQ ID NO: 302) and pCLS15794 (SEQ ID NO: 303), respectively. To generate an intact coding sequence by *in vivo* homologous recombination, approximately 40 ng of each plasmid linearized by digestion with BssHII and 1 ng of the pCLS0542 (SEQ ID NO: 156) plasmid DNA linearized by digestion with NcoI and EagI were used to transform, respectively, the yeast *S. cerevisiae* strain FYC2-6A (*MATα*, *trp1*Δ63, *leu2*Δ1, *his3*Δ200) using a high efficiency LiAc transformation protocol (Arnould et al. 2007).

All the yeast target reporter plasmids containing the TALEN or meganuclease DNA target sequences were constructed as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

The TALE::scTB2aD01-based constructs were tested in a yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006) on hybrid targets TCRB02Tsp7 (SEQID NO: AC4), TCRB02Tsp12 (SEQID NO: AC5) and TCRB02Tsp16 (SEQID NO: AC6), illustrated in figure 14. The TCRB02.1-only target was included to compare activity with the engineered TCRB02-A meganuclease (pCLS6857, SEQ ID NO: 291), which does not require the TALE DNA binding sites for activity. Activity levels on the respective targets in yeast cells for the indicated TALE::scTB2aD01-based constructs are shown in figure 14. Notably, under the *in* vivo conditions tested the TALE-TB2A2::scTB2aD01_148C (pCLS15794, SEQ ID NO: 303) construct no longer cleaves targets lacking the DNA sequence recognized by the TALE DNA binding moiety.

Activity of TALE::CreI in mammalian cells

DNA encoding the TALE-TB2A2::scTB2aD01 and TALE-TB2A3::scTB2aD01 constructs from pCLS15792 (SEQ ID NO: 301) and pCLS15793 (SEQ ID NO: 302) were subcloned into the pCLS1853 (SEQ ID NO: 193) mammalian expression plasmid using AscI and XhoI restriction enzymes for the receiving plasmid and BssHII and XhoI restriction enzymes for TALE::scTB2aD01-based inserts, leading to the mammalian expression plasmids pCLS14894 and pCLS14895 (SEQ ID NO: 308 and 309), respectively.

All mammalian target reporter plasmids containing the TALEN DNA target sequences were constructed using the standard Gateway protocol (INVITROGEN) into a CHO reporter vector (Arnould, Chames et al. 2006, Grizot, Epinat et al. 2010).

To monitor protein expression levels, TALE::scTB2aD01-based constructs were transfected in mammalian cells (HEK293) alongside the engineered TCRB02-A meganuclease (pCLS6857, SEQ ID NO: 291). Briefly, cells were transfected, respectively, with 300 ng of each protein encoding plasmid in the presence of lipofectamine. Fourty-eight hours post-transfection, 20 µg of total protein extract for each sample was analyzed by Western-Blot using a polyclonal anti-I-CreI antibody. A typical western-blot is shown in Figure 15.

Relative toxicity of the TALE::scTB2aD01-based constructs was assessed using a cell survival assay. CHOK1 cells were used to seed plates at a density of $2.5*10^3$ cells per well. The following day, varying amounts of plasmid encoding either the TALE::scTB2aD01-based constructs (pCLS14894 and pCLS14895; SEQ ID NO: 308 and 309) or the engineered TCRB02-A meganuclease (pCLS6857, SEQ ID NO: 291) and a constant amount of GFP-encoding plasmid (10 ng) were used to transfect the cells with a total quantity of 200 ng using Polyfect reagent. GFP levels were monitored by flow cytometry (Guava Easycyte, Guava technologies) on days 1 and 6 post-transfection. Cell survival is expressed as a percentage, calculated as a ratio (TALEN and meganuclease-transfected cells expressing GFP on Day 6 / control-transfected cells expressing GFP on Day 6) corrected for the transfection efficiency determined on Day 1. Typical cell survival assay data are shown in Figure 16.

Cleavage activity *in* vivo was monitored via detection of NHEJ events in the presence of TREX2 exonuclease. Plasmid (3 µg) encoding either the TALE::scTB2aD01-based constructs (pCLS14894 and pCLS14895; SEQ ID NO: 308 and 309) or the engineered TCRB02-A meganuclease (pCLS6857, SEQ ID NO: 291) and 2 µg of scTrex2-encoding plasmid (pCLS8982, SEQ ID NO: 310) were used to transfect the HEK293 cells in the presence of lipofectamine. Genomic DNA was extracted 2 and 7 days post-transfection with the DNeasy Blood and Tissue kit (Qiagen) and the region encompassing the TCRB02 site (Figure 13) was amplified using the PCR with oligos TRBC2F3 (Seq ID NO: 311) and TRBC2R3B (SEQ ID NO: 312) at day 2 post-transfection and with oligos TRBC2F4 (SEQID NO: 315) and TRBC2R4B (SEQID NO: 314) at day 7 post-transfection. Respective PCR products (100 ng) were heat denatured, allowed to re-anneal by slow-cooling then treated with T7 endonuclease 1 (NEB) for 15 minutes at 37°C. Digested PCR products are separated on 10% acrylamide gels and visualized with SYBRgreen (Invitrogen) staining. Cleavage of mismatched DNA sequences by T7 endonuclease is indicative of NHEJ events resulting from the activity of the cTALEN or meganuclease at the targeted locus. Figure 17 illustrates the detectable NHEJ activity of the TALE::scTB2aD01-based constructs (pCLS14894 and pCLS14895; SEQ ID NO: 308 and 309) compared to the engineered

TCRB02-A meganuclease (pCLS6857, SEQ ID NO: 291). Whereas at day 2 NHEJ results are comparable for all constructs, NHEJ activity at day 7 can only be detected for the TALE::scTB2aD01-based constructs., suggesting that these compact TALENs do not induce cytotoxicity.

Engineering of the TALE::CreI

A significant novel property of the TALE::CreI compact TALEN resides in the ability to independently engineer the "hybrid" specificity of the final molecule. As such, the inherent activity/specificity ratio can be modulated within the TALE::CreI-derived constructs, allowing for unprecedented specific targeting with retention of high DNA cleavage activity. In its simplest form, successful re-targeting of the TALE DNA binding domain is achieved via the RVD cipher (Figure 3), with a pseudo one-to-one correspondence to the underlying DNA base. Engineering of the I-CreI moiety, however, presents more challenges insomuch as there exists a potential codependence of protein-DNA contacts needed for suitable DNA binding and cleavage activity. Methods have been described (WO2006097854, WO2008093249, WO03078619, WO2009095793, WO 2007/049095, WO 2007/057781, WO 2006/097784, WO 2006/097853, WO 2007/060495, WO 2007/049156 and WO 2004/067736) to successfully re-engineer the I-CreI meganuclease to target novel DNA sequences. As some of these methods rely on a clustered approach, it can be envisioned that using said approach the "absolute" specificity of the I-CreI moiety could be reduced in a stepwise manner. For example, the breakdown of the I-CreI DNA interaction surface into discrete 10NNN, 7NN, 5NNN and 2NN regions (per monomeric subunit half) allows for novel engineering wherein high specificity is maintained in the central 5NNN-2NN region at the expense of "loose" or broad specificity in the outer 10NNN-7NN regions. In essence such an approach could reduce the complexity of re-engineering the I-CreI-derived scaffold for a compact TALEN context as only "selectivity" in cleavage is required for the catalytic domain, with subsequent specificity provided by the TALE DNA binding part of the protein fusion. Taken together, the ease of engineering combined with the potential high specificity and high DNA cleavage activity make TALE::CreI-derived compact TALENs ideal tools for therapeutic applications. Finally, it should be noted that the I-CreI moiety could in principle be replaced with a host of naturally existing or re-engineered homing endonuclease-derived catalytic domains.

Example 3f: Activity of TALE::SnaseSTAUU

Activity of TALE::SnaseSTAUU in yeast

Variants differing by truncations of the C-terminal domain of the AvrBs3-derived TALEN (SEQ ID NO: 196) are chosen as starting scaffolds. A subset of these variants includes truncation after positions G914 (C28) and L926 (C40) (the protein domains of truncated C-terminal domains C28 and C40 are respectively given in SEQ ID NO: 205 and 206). The plasmids coding for the variant scaffolds containing the AvrBs3-derived N-terminal domain, the AvrBs3-derived set of repeat domains and the truncated AvrBs3-derived C-terminal domain [pCLS7807 and pCLS7809, (SEQ ID NO: 213 and 214) which are based on the pCLS7184 (SEQ ID NO: 196)] allow cloning of any catalytic domain in fusion to the C-terminal domain, using the restriction sites BamHI and EagI.

The DNA corresponding to amino acid residues 83 to 231 of SnaseSTAAU (SEQ ID NO: 30) is amplified by the PCR to introduce, at the DNA level, a BamHI (at the 5' of the coding strand) and a EagI (at the 3' of the coding strand) restriction site and, at the protein level, a linker (for example -SGGSGS- stretch, SEQ ID NO: 219) between the C terminal domain of the TALE and the SnaseSTAAU catalytic domain. The final TALE::SnaseSTAAU constructs are generated by insertion of the SnaseSTAAU catalytic domain into the scaffold variants using BamHI and EagI and standard molecular biology procedures. Scaffold variants truncated after positions G914 (C28) and L926 (C40), respectively encoded by pCLS7807 and pCLS7809, (SEQ ID NO: 213 and 214), were fused to the SnaseSTAAU catalytic domain (SEQ ID NO: 30), leading to pCLS9082 and pCLS9081 (SEQ ID NO: 370 and 371). The cloning step also brings at the amino acid level an AAD sequence at the Cter of the SnaseSTAAU catalytic domain.

All the yeast target reporter plasmids containing the TALEN DNA target sequences were constructed as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

The TALE-AvrBs3:: SnaseSTAAU constructs were tested in a yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006) on pseudo palindromic targets in order to compare activity with a standard TALE-AvrBs3::FokI TALEN, which requires two binding sites for activity. AvrBs3 targets contain two identical recognition sequences juxtaposed with the 3' ends proximal and separated by "spacer" DNA ranging from 5 to 40 bps (SEQ ID NO: 157 to 192, Table 7). In addition, TALE-AvrBs3::SnaseSTAAU constructs were tested on a target having only a single AvrBs3 recognition site (SEQ ID NO: 224). Data summarized in figure 19 show that TALE-AvrBs3::SnaseSTAAU constructs are active on targets having two AvrBs3 recognition site, according to the chimeric protein of the present invention, but also on targets containing only one AvrBs3 recognition site.

Example 4

Basic cTALENs are composed of a single DNA binding domain fused to a single catalytic domain and are designed to stimulate HR via a single double-strand DNA cleavage or single-strand nicking event. For certain applications (e.g. gene inactivation), it is favorable to enhance the level of NHEJ. This example illustrates the creation of a dual-cleavage cTALEN (dcTALEN) that is capable of effecting cleavage of double-strand DNA at two distinct sites flanking the TALE DNA binding domain (Figure 5C). The simultaneous cleavage of the DNA at the two sites is expected to eliminate the intervening sequence and therefore abolish "scarless" re-ligation by NHEJ (Figure1).

The baseline scaffolds (SEQ ID NO: 136 to SEQ ID NO: 139) described in Example 3 are used as starting points for fusion designs. A non-exhaustive list of catalytic domains amenable to fusion with TALE DNA binding domains is presented in Table 2. A non-exhaustive list of linkers that can be used is presented in Table 3. See examples 3, 5, 6 and 7 for additional details concerning the choice of linker or enhancement domain. For the dcTALEN designs, at least one cleavase domain is fused (N- or C-terminal) to the TALE DNA binding domain. The additional catalytic domain can be either a nickase of cleavase (endonuclease or exonuclease) domain, and depends on the nature of the application. For example, the coupling of a cleavase domain on one side with a nickase domain on the other could result in excision of a single-strand of DNA spanning the TALE DNA binding region. The targeted generation of extended single-strand overhangs could be applied in applications that target DNA repair mechanisms. For targeted gene inactivation, the use of two cleavase domains in the dcTALEN is preferred.

All dcTALEN designs are assessed using our yeast assay (see Example 1) and provide detectable activity comparable to existing engineered meganucleases. Furthermore, potential enhancements in NHEJ are monitored using the mammalian cell based assay as described in Example 3.

Example 4a: Activity of TevI::TALE::FokI and TevI::TALE::TevI dual cleavage TALENs

Dual cleavage TALENs (CD::TALE::CD), possessing an N-terminal I-TevI-derived catalytic domain and a C-terminal catalytic domain derived from either FokI (SEQ ID NO :368) or I-TevI (SEQ ID NO: 20), were generated on the baseline bT2-Avr (SEQ ID NO: 137) scaffold. The catalytic domain fragment of I-TevI was excised from plasmid pCLS12731 (SEQ ID NO: 236) and subcloned into vectors pCLS15795

(SEQ ID NO: 351) and pCLS9013 (SEQ ID NO: 153) by restriction and ligation using NcoI and NsiI restriction sites, yielding TevD02_cT11Avr_FokI-L (pCLS15796, SEQ ID NO: 352, encoding the protein of SEQ ID NO: 447) and TevD02_cT11Avr_TevD02 (pCLS15797, SEQ ID NO: 353, encoding the protein of SEQ ID NO: 448), respectively. All constructs were sequenced and the insert transferred to additional vectors as needed (see below).

The final TevI::TALE-AvrBs3::FokI and TevI::TALE-AvrBs3::TevI yeast expression plasmids, pCLS13299 (SEQ ID NO: 354, encoding the protein of SEQ ID NO: 449) and pCLS13301 (SEQ ID NO: 355, encoding the protein of SEQ ID NO: 450), were prepared by yeast *in vivo* cloning using plasmids pCLS15796 (SEQ ID NO: 352) and pCLS15797 (SEQ ID NO: 353), respectively. To generate an intact coding sequence by *in vivo* homologous recombination, approximately 40 ng of each plasmid linearized by digestion with BssHII and 1 ng of the pCLS0542 (SEQ ID NO: 156) plasmid DNA linearized by digestion with NcoI and EagI were used to transform, respectively, the yeast *S. cerevisiae* strain FYC2-6A (*MATα*, *trp1*Δ63, *leu2*Δ1, *his3*Δ200) using a high efficiency LiAc transformation protocol (Arnould et al. 2007).

All the yeast target reporter plasmids containing the TALEN DNA target sequences were constructed as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

The TevI::TALE-AvrBs3::FokI and TevI::TALE-AvrBs3::TevI constructs were tested in a yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006) on pseudo palindromic targets in order to compare activity with a standard TALE-AvrBs3::FokI TALEN (pCLS8590, SEQ ID NO: 244), which requires two binding sites for activity. AvrBs3 targets contain two identical recognition sequences juxtaposed with the 3' ends proximal and separated by "spacer" DNA ranging from 5 to 40 bps (SEQ ID NO: 157 to 192, Table 6). In addition, constructs were tested on a target having only a single AvrBs3 or RagT2-R recognition site (SEQ ID NO: 238, Table 11). On suitable targets, the TevI::TALE-AvrBs3::FokI and TevI::TALE-AvrBs3::TevI activity levels in yeast were comparable to those of their parent molecules lacking the N-terminal I-TevI-derived catalytic domain. Significant activity is illustrated in table 11 for a sample single-site target, according to the dcTALEN of the present invention.

	Target DNA	
TALEN Construct	Avr25 (dual-site) (SEQ ID NO: 177)	Avr25RAGT2R (single-site)(SEQ ID NO: 238)
TevI::TALE-AvrBs3	++++	+++

TALE-AvrBs3::FokI	++++	n.d.
TALE-AvrBs3::TevI	++++	n.d.
TevI::TALE-AvrBs3::FokI	++++	+++
TevI::TALE-AvrBs3::TevI	++++	+++

Table 11: Activity of various cTALENs and dcTALENs on dual- and single-site DNA targets. Relative activity is scaled as: n.d., no activity detectable; +, <25% activity; ++, 25% to <50% activity; +++, 50% to <75% activity; ++++, 75% to 100% activity.

Example 4b: scTrex2::TALE::FokI dual cleavage TALEN

A dual cleavage TALEN (CD::TALE::CD), possessing an N-terminal scTrex2-derived catalytic domain and a C-terminal catalytic domain derived from FokI, was generated on the baseline bT2-Avr (SEQ ID NO: 137) scaffold. The catalytic domain fragment of scTrex2 was excised from plasmid pCLS15798 (SEQ ID NO: 356, encoding the protein of SEQ ID NO: 451) and subcloned into vector pCLS15795 (SEQ ID NO: 351) by restriction and ligation using NcoI and NsiI restriction sites, yielding scTrex2_cT11Avr_FokI-L (pCLS15799, SEQ ID NO: 357, encoding the protein of SEQ ID NO: 452). The construct was sequenced and the insert transferred to additional vectors as needed (see below).

DNA encoding the TALE-AvrBs3::FokI or scTrex2::TALE-AvrBs3::FokI constructs from either pCLS15795 (SEQ ID NO: 351) or pCLS15799 (SEQ ID NO: 357), respectively, was subcloned into the pCLS1853 (SEQ ID NO: 193) mammalian expression plasmid using AscI and XhoI restriction enzymes for the receiving plasmid and BssHII and XhoI restriction enzymes for the inserts, leading to the mammalian expression plasmids pCLS14972 and pCLS14971 (SEQ ID NO: 358 and 359), respectively.

All mammalian target reporter plasmids containing the TALEN DNA target sequences were constructed using the standard Gateway protocol (INVITROGEN) into a CHO reporter vector (Arnould, Chames et al. 2006, Grizot, Epinat et al. 2010). The TALE-AvrBs3::FokI and scTrex2::TALE-AvrBs3::FokI constructs were tested in an extrachromosomal assay in mammalian cells (CHO K1) on pseudo palindromic targets in order to compare activity with a standard TALE-AvrBs3::FokI TALEN, which requires two binding sites for activity. AvrBs3 targets contain two identical recognition sequences juxtaposed with the 3' ends proximal and separated by "spacer" DNA ranging from 5 to 40 bps (SEQ ID NO: 157 to 192, Table 6).

For this assay, CHO K1 cells were transfected in a 96-well plate format with 75 ng of target vector and an increasing quantity of each variant DNA from 0.7 to 25 ng , in the presence of PolyFect reagent (1µL per well). The total amount of transfected DNA was completed to 125 ng (target DNA, variant DNA, carrier DNA) using an empty vector. Seventy-two hours after transfection, culture medium was removed and 150µl of lysis/revelation buffer for β-galactosidase liquid assay

was added. After incubation at 37°C, optical density was measured at 420 nm. The entire process is performed on an automated Velocity11 BioCel platform (Grizot, Epinat et al. 2009).

Activity levels in mammalian cells on suitable targets for the scTrex2::TALE-AvrBs3::FokI construct were comparable to those of the parent TALE-AvrBs3::FokI molecule, indicating that the extra scTrex2 moiety does not impair the TALEN DNA cleavage function. Assessment of the scTrex2 function is performed in assays suitable for the detection of NHEJ events.

Example 5

Baseline designs for the cTALEN scaffolds are based on established TALE DNA binding domains. Compact TALENs are designed to be as small and efficient as possible. To obtain this goal it may therefore be necessary to enlist "enhancer" domains to bridge the functional gap between compact TALE DNA binding domains and the various catalytic domains. Figure 6 (A-E) illustrates various non-exhaustive configurations wherein such enhancer domains can be applied. Note that the figure is illustrative only, and N- vs. C-terminal variations are implied (i.e. Figure 6A can also have an N-terminal enhancer domain and C-terminal catalytic domain). Tables 1 and 2 lists potential enhancer domains that could assist in DNA binding (specific and non-specific contacts).

Enhanced TALENs (eTALENs) are created using functional cTALENS from Example 3. The addition of the enhancer domain is evaluated in our yeast assay (see Example 1). A particular enhancer domain is judged useful if it provides a minimal 5% enhancement in efficiency of the starting cTALEN, more preferably a minimal 10% enhancement, more preferably 20%, more preferably 30%, more preferably 40%, more preferably 50%, again more preferably an enhancement greater than 50%.

Example 5a: TALE::ColE7::TALE enhanced TALENs

Enhanced TALENs (TALE::CD::TALE), possessing N- and C-terminal TALE DNA binding domains bordering a central DNA cleavage domain, were generated using the sT2 (SEQ ID NO: 135) core scaffold. The layout of this class of compact TALEN is illustrated in figure 6B, wherein the N-terminal "enhancer domain" is itself a TALE DNA binding domain. A point-mutant derivative of the ColE7 catalytic domain (pCLS15785, SEQ ID NO: 285) was chosen for the catalytic core of the eTALEN. Two final constructs, TALE-AvrBs3::ColE7_A497::TALE-RagT2-R (pCLS15800, SEQ ID NO: 360, encoding the protein of SEQ ID NO: 453) and TALE-RagT2-R::ColE7_A497::TALE-AvrBs3 (pCLS15801, SEQ ID NO: 361, encoding the protein of SEQ ID NO: 454), were obtained using standard molecular cloning

techniques with DNA sequences from sT2 (SEQ ID NO: 135), pCLS15785 (SEQ ID NO: 285), AvrBs3 (SEQ ID NO: 152) and RagT2-R (SEQ ID NO: 271) as templates. All TALE::CD::TALE constructs were sequenced and the inserts transferred to additional vectors as needed (see below).

The final TALE::CD::TALE-based yeast expression plasmids, pCLS12106 (SEQ ID NO: 362) and pCLS12110 (SEQ ID NO: 363, were prepared by restriction and ligation using NcoI and EagI restriction sites to subclone into the pCLS0542 (SEQ ID NO: 156) plasmid. The yeast *S. cerevisiae* strain FYC2-6A (*MAT*α, *trp1*Δ63, *leu2*Δ1, *his3*Δ200) was transformed using a high efficiency LiAc transformation protocol (Arnould et al. 2007).

All the yeast target reporter plasmids containing the TALEN DNA target sequences were constructed as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

The TALE::CD::TALE constructs are tested in a yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006) on asymmetric AvrBs3/RagT2-R hybrid targets in order to compare activity with a parent compact TALEN (e.g. pCLS8589, SEQ ID NO: 233), which has activity on targets with a single binding site. In addition, constructs are tested on a target having only a single AvrBs3 or RagT2-R recognition site.

Example 6

To date, all known TAL effectors and derivatives thereof appear to require a T base at positions -1 (Figure 3) in the recognition sequence. To overcome this limitation, an enhancer domain is used to replace the N-terminal region of the TALE protein. Sequence and structure-based homology modeling of the N-terminal TALE region of bT2-derivatives have yielded three potential candidate proteins (Table 1): (i) Fem-3 binding factor, SEQ ID NO: 4,(FBF1, Puf family of RNA binding proteins) from C.elegans; (ii) artificial alpha-helicoidal repeat proteins (αRep), SEQ ID NO: 5 and; (iii) proteins of the Ankyrin super- family. The content and arrangement of secondary structure elements allows for using these models as starting points for enhancer domains that replace the N-terminal region of the TALE protein.

Chimeric proteins are constructed using the analogous regions from one of the 3 candidates mentioned to replace the N-terminal TALE protein region up to the first canonical repeat domain. The new interface is redesigned *in silico*, using the homology models as guides. This approach can be

used to pinpoint the determinants of specificity for the requisite T at position -1 of the target sequence. The replacement enhancer domain should at minimum provide structural integrity to the cTALEN protein. Constructs are evaluated in our yeast assay (see Example 1). A particular enhancer domain is judged useful if it provides a minimal 5% retention in activity of the starting cTALEN in the absence of a T at target position -1, more preferably a minimal 10% retention, more preferably 20%, more preferably 30%, more preferably 40%, more preferably 50%, again more preferably a retention in activity greater than 50%.

Example 7

To generate more suitable and compact scaffolds for cTALENS, the nature of the C-terminal region (beyond the final half-repeat domain) of the TALE protein has been analyzed. Sequence and structure-based homology modeling of the C-terminal TALE region of bT2-derivatives have yielded three potential candidate proteins (Table 1): (i) the hydrolase/transferase of Pseudomonas Aeuriginosa, SEQ ID NO: 6 ; (ii) the Polymerase domain from the Mycobacterium tuberculosis Ligase D, SEQ ID NO: 7; (iii) initiation factor eIF2 from Pyrococcus, SEQ ID NO: 8; (iv) Translation Initiation Factor Aif2betagamma, SEQ ID NO: 9. As in example 6, homology models are used to pinpoint regions for generating possible C-terminal truncations; potential truncation positions include 28, 40, 64, 118, 136, 169, 190 residues remaining beyond the last half-repeat domain. Additionally, homologous regions from the aforementioned proteins can be used to replace the C-terminal domain entirely. Contact prediction programs can be used to identify, starting from the primary sequence of a protein, the pairs of residues that are likely proximal in the 3D space. Such chimeric proteins should provide more stable scaffolds on which to build cTALENs.

Constructs are evaluated in our yeast assay (see Example 1). A particular enhancer domain is judged useful if it provides a minimal 5% retention in activity of the starting cTALEN, more preferably a minimal 10% retention, more preferably 20%, more preferably 30%, more preferably 40%, more preferably 50%, again more preferably a retention in activity greater than 50%.

Example 8

To generate compact TALENS with alternative activities, trans cTALENS are generated by (a) using a catalytic domain with separable activities (Figure 7A, B), or; (b) providing an auxiliary activity as a TALE-fusion (Figure 7C). Sequence and structure-based modeling of class III (Chan, Stoddard et al.

2011). TypeIIS restriction endonucleases (REases) were used to create trans TALENs (see Table 2 for a non-exhaustive list). The initial trans TALEN is generated via fusion of an independently active catalytic domain (e.g. the Nt.BspD6I nickase) as described in Examples 3 and 4. In principle this trans TALEN can be used as is depending on the application. To convert the cTALEN to a functional trans TALEN, the auxiliary domain (in this case, ss.BspD6I) is provided *in trans* (Figure 8A). Such optionally trans and/or heterodimeric proteins can allow for cTALEN scaffolds with activity that can be modulated to a given application.

Constructs are evaluated in our yeast assay (see Example 1). A particular auxiliary domain is judged useful if it provides an alternative activity to that of the starting cTALEN.

If the auxiliary domain used exhibits activity independent of the initial cTALEN (i.e in a non-trans TALEN context), it can as well be fused to a TALE domain for specific targeting (Figure 7B). Auxiliary domains can also be provided *in trans* as targeted entities to provide functions unrelated to the cTALEN (Figure 7C).

Example 8a: Specific inhibition of TALEN catalytic activity

As mentioned in examples 3c and 3d, both NucA (SEQ ID NO: 26) and ColE7 (SEQ ID NO: 140) can be inhibited by complex formation with their respective inhibitor proteins, NuiA (SEQ ID NO: 229) and Im7 (SEQ ID NO: 230). Colicin-E9 (SEQ ID NO: 366) is another non-limiting example of protein which can be inhibited by its respective inhibitor Im9 (SEQ ID NO: 369). With respect to TALENs derived from the NucA (TALE::NucA) or ColE7 (TALE::ColE7) catalytic domains, the inhibitors serve as auxiliary domains (Figure 7A) that modulate the activity by preventing DNA cleavage.

The Im7 (SEQ ID NO: 230) and NuiA (SEQ ID NO: 229) inhibitor proteins were subcloned into the pCLS7763 backbone (SEQ ID NO: 241) by restriction and ligation using NcoI and EagI restriction sites, yielding pCLS9922 (SEQ ID NO: 242) and pCLS9923 (SEQ ID NO: 243), respectively. These plasmids were then used in co-transformation experiments in the standard yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

All the yeast target reporter plasmids containing the TALEN DNA target sequences were constructed as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

TALE-AvrBs3::NucA (pCLS9924, SEQ ID NO: 223) and TALE-AvrBs3::ColE7 (pCLS8589, SEQ ID NO: 233) constructs were tested in a yeast SSA assay on pseudo palindromic targets in order to compare activity with a standard TALE-AvrBs3::FokI TALEN (pCLS8590, SEQ ID NO: 244), which requires two binding sites for activity. AvrBS3 targets contain two identical recognition sequences juxtaposed with the 3' ends proximal and separated by "spacer" DNA ranging from 5 to 40 bps (SEQ ID NO: 157 to 192, Table 7). In addition, constructs were tested on a target having only a single AvrBs3 recognition site (SEQ ID NO: 224, Table 7). Activity modulation of the TALENs was assessed in the presence or absence of specific or unspecific inhibitor protein, using the TALE-AvrBs3::FokI TALEN as control.

Data summarized in table 12 indicate that TALE-AvrBs3::NucA and TALE-AvrBs3::ColE7 constructs are specifically inactivated by the presence of their respective inhibitor proteins NuiA and Im7, according to the present invention.

	Inhibitor Protein		
TALEN Construct	None	NuiA	Im7
TALE-AvrBs3::NucA (SEQ ID NO: 223)	++++	n.d.	++++
TALE-AvrBs3::ColE7 (SEQ ID NO: 233)	++++	++++	n.d.
TALE-AvrBs3::FokI (SEQ ID NO: 244)	++++	++++	++++

Table 12: Activity of TALEN constructs in the presence of inhibitor protein. Relative activity is scaled as: n.d., no activity detectable; +, <25% activity; ++, 25% to <50% activity; +++, 50% to <75% activity; ++++, 75% to 100% activity.

Example 8b: Enhancing TALEN catalytic activity via a trans TALEN

Example 3b illustrates that the TevI::TALE functions unassisted as a compact TALEN (pCLS8522, SEQ ID NO: 237). To further enhance activity, a trans TALEN was designed using a TALE::TevI construct in a layout depicted in figure 7C. The DNA sequence coding for the RVDs to target the RagT2-R site (SEQ ID NO: 271) was subcloned into plasmid pCLS7865-cT11_TevD02 (pCLS9011, SEQ ID NO: 151) using Type IIS restriction enzymes BsmBI for the receiving plasmid and BbvI and SfaNI for the inserted RVD sequence to create the subsequent TALE-RagT2-R::TevI construct cT11RagT2-R_TevD02 (pCLS15802, SEQ ID NO: 364). The construct was sequenced and the insert subcloned into the pCLS7763 backbone (SEQ ID NO: 241) by restriction and ligation using NcoI and EagI restriction sites, yielding pCLS8990 (SEQ ID NO: 365). Plasmid pairs pCLS8522 (SEQ ID NO: 237) and pCLS7763 (SEQ ID

NO: 241) or pCLS8522 (SEQ ID NO: 237) and pCLS8990 (SEQ ID NO: 365) were then used in co-transformation experiments in the standard yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006).

All the yeast target reporter plasmids containing the TALEN DNA target sequences were constructed as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006). The TALE-RagT2-R::TevI/TevI::TALE-AvrBs3 construct pairs were tested in a yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006) on asymmetric RagT2-R /AvrBs3 hybrid targets in order to compare activity with a parent compact TALEN (e.g. pCLS8522, SEQ ID NO: 237), which has activity on targets with a single binding site. RagT2-R/AvrBs3 hybrid targets contain two different recognition sequences juxtaposed with the 3' end of the first (RagT2-R) proximal to the 5' end of the second (AvrBs3) and separated by "spacer" DNA ranging from 5 to 40 bps (SEQ ID NO: G064 to G099, Table 13). Figure 18 illustrates the modulation in TevI::TALE-AvrBs3 activity provided by the TALE-RagT2-R::TevI construct, according to the trans cTALEN of the present invention.

Example 9: Replacement of the C-terminal domain by a polypeptide linker, activity with colE7 catalytic domain.

We generated a first library of 37 different linkers. Many of them have a common structure comprising a variable region encoding 3 to 28 amino acids residues and flanked by regions encoding SGGSGS stretch (SEQ ID NO: 219) at both the 5' and a 3' end (SEQ ID NO: 372 to 408). These linkers contain XmaI and BamHI restriction sites in their 5' and 3' ends respectively. The linker library is then subcloned in pCLS7183 (SEQ ID NO: 141) via the XmaI and BamHI restriction sites to replace the C-terminal domain of the AvrBs3-derived TALEN (pCLS7184, SEQ ID NO: 196). The AvrBs3-derived set of repeat domains (RVDs) or any other RVD sequences having or lacking the terminal half RVD is cloned in this backbone library. DNA from the library is obtained, after scrapping of the colonies from the Petri dishes, using standard miniprep techniques. The FokI catalytic head is removed using BamHI and EagI restriction enzymes, the remaining backbone being purified using standard gel extraction techniques.

DNA coding for ColE7 catalytic domain (SEQ ID NO: 11) was amplified by the PCR to introduce, at the DNA level, a BamHI (at the 5' of the coding strand) and a EagI (at the 3' of the coding strand) restriction site and, at the protein level, a linker (for example -SGGSGS - stretch, SEQ ID NO: 219)

between the C terminal domain library and the catalytic head. After BamHI and EagI digestion and purification, the DNA coding for the different catalytic heads were individually subcloned into the library scaffold previously prepared.

DNA from the final library is obtained, after scrapping of the colonies from Petri dishes, using standard miniprep techniques and the resulting libraries are screened in our yeast SSA assay as previously described (International PCT Applications WO 2004/067736 and in Epinat, Arnould et al. 2003; Chames, Epinat et al. 2005; Arnould, Chames et al. 2006; Smith, Grizot et al. 2006) on pseudo palindromic targets in order to compare activity with a standard TALE-AvrBs3::FokI TALEN, which requires two binding sites for activity. AvrBs3 targets contain two identical recognition sequences juxtaposed with the 3' ends proximal and separated by "spacer" DNA containing 15, 18, 21 and 24 bps (SEQ ID NO: 167, 170, 173 and 176, Table 7). In addition, constructs (SEQ ID NO: 416-419) were tested on a target having only a single AvrBs3 recognition site (SEQ ID NO: 224). Data summarized in figure 20 show sequences of the linker of a fraction of ColE7 constructs being active on targets having two AvrBs3 recognition sites or only one AvrBs3 recognition site.

List of cited references

Arimondo, P. B., C. J. Thomas, et al. (2006). "Exploring the cellular activity of camptothecin-triple-helix-forming oligonucleotide conjugates." Mol Cell Biol **26**(1): 324-33.

Arnould, S., P. Chames, et al. (2006). Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. Journal of Molecular Biology. **355:** 443-58.

Arnould, S., P. Chames, et al. (2006). "Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets." J Mol Biol **355**(3): 443-58.

Arnould, S., C. Delenda, et al. (2011). "The I-CreI meganuclease and its engineered derivatives: applications from cell modification to gene therapy
gzq083 [pii] 10.1093/protein/gzq083." Protein engineering, design & selection : PEDS **24**(1-2): 27-31.

Arnould, S., C. Perez, et al. (2007). "Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells." Journal of Molecular Biology **371**(1): 49-65.

Ashworth, J., J. J. Havranek, et al. (2006). "Computational redesign of endonuclease DNA binding and cleavage specificity." Nature **441**(7093): 656-9.

Bedayat, B., A. Abdolmohamadi, et al. (2010). "Sequence-specific correction of genomic hypoxanthine-guanine phosphoribosyl transferase mutations in lymphoblasts by small fragment homologous replacement
10.1089/oli.2009.0205." Oligonucleotides **20**(1): 7-16.

Bennardo, N., A. Cheng, et al. (2008). "Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair." PLoS Genet **4**(6): e1000110.

Bennardo, N., A. Gunn, et al. (2009). "Limiting the persistence of a chromosome break diminishes its mutagenic potential." PLoS Genet **5**(10): e1000683.

Boch, J., H. Scholze, et al. (2009). "Breaking the code of DNA binding specificity of TAL-type III effectors." Science **326**(5959): 1509-12.

Boch, J., H. Scholze, et al. (2009). "Breaking the code of DNA binding specificity of TAL-type III effectors
1178811 [pii] 10.1126/science.1178811." Science **326**(5959): 1509-12.

Bolduc, J. M., P. C. Spiegel, et al. (2003). "Structural and biochemical analyses of DNA and RNA binding by a bifunctional homing endonuclease and group I intron splicing factor." Genes Dev **17**(23): 2875-88.

Buis, J., Y. Wu, et al. (2008). "Mre11 nuclease activity has essential roles in DNA repair and genomic stability distinct from ATM activation." Cell **135**(1): 85-96.

Capecchi, M. R. (2001). "Generating mice with targeted mutations 10.1038/nm1001-1086 nm1001-1086 [pii]." Nature Medicine **7**(10): 1086-90.

Carroll, D. (2008). "Progress and prospects: zinc-finger nucleases as gene therapy agents gt2008145 [pii] 10.1038/gt.2008.145." Gene therapy **15**(22): 1463-8.

Chames, P., J. C. Epinat, et al. (2005). "In vivo selection of engineered homing endonucleases using double-strand break induced homologous recombination." Nucleic Acids Res **33**(20): e178.

Chames, P., J. C. Epinat, et al. (2005). "In vivo selection of engineered homing endonucleases using double-strand break induced homologous recombination." Nucleic Acids Research **33**(20): e178.

Chan, S. H., B. L. Stoddard, et al. (2011). "Natural and engineered nicking endonucleases--from cleavage mechanism to engineering of strand-specificity." Nucleic Acids Research **39**: 1-18.

Chevalier, B., M. Turmel, et al. (2003). "Flexible DNA target site recognition by divergent homing endonuclease isoschizomers I-CreI and I-MsoI." J Mol Biol **329**(2): 253-69.

Chevalier, B. S., T. Kortemme, et al. (2002). "Design, activity, and structure of a highly specific artificial endonuclease." Mol Cell **10**(4): 895-905.

Chevalier, B. S., R. J. Monnat, Jr., et al. (2001). "The homing endonuclease I-CreI uses three metals, one of which is shared between the two active sites." Nat Struct Biol **8**(4): 312-6.

Chevalier, B. S. and B. L. Stoddard (2001). "Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility." Nucleic Acids Res **29**(18): 3757-74.

Choo, Y. and A. Klug (1994). "Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions." Proceedings of the National Academy of Sciences of the United States of America **91**(23): 11168-72.

Choo, Y. and A. Klug (1994). "Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage." Proceedings of the National Academy of Sciences of the United States of America **91**(23): 11163-7.

Choulika, A., A. Perrin, et al. (1995). "Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae." Mol Cell Biol **15**(4): 1968-73.

Christian, M., T. Cermak, et al. (2010). "Targeting DNA double-strand breaks with TAL effector nucleases." Genetics **186**(2): 757-61.

Christian, M., T. Cermak, et al. (2010). "Targeting DNA double-strand breaks with TAL effector nucleases genetics.110.120717 [pii] 10.1534/genetics.110.120717." Genetics **186**(2): 757-61.

Cost, G. J., Y. Freyvert, et al. (2010). "BAK and BAX deletion using zinc-finger nucleases yields apoptosis-resistant CHO cells 10.1002/bit.22541." Biotechnology and Bioengineering **105**(2): 330-40.

Delacote, F. and B. S. Lopez (2008). "Importance of the cell cycle phase for the choice of the appropriate DSB repair pathway, for genome stability maintenance: the trans-S double-strand break repair model 5149 [pii]." Cell Cycle **7**(1): 33-8.

Doudeva, L. G., H. Huang, et al. (2006). "Crystal structural analysis and metal-dependent stability and activity studies of the ColE7 endonuclease domain in complex with DNA/Zn2+ or inhibitor/Ni2+ 15/2/269 [pii] 10.1110/ps.051903406." Protein science : a publication of the Protein Society **15**(2): 269-80.

Doyon, J. B., V. Pattanayak, et al. (2006). "Directed evolution and substrate specificity profile of homing endonuclease I-SceI." Journal of the American Chemical Society **128**(7): 2477-84.
Doyon, J. B., V. Pattanayak, et al. (2006). "Directed evolution and substrate specificity profile of homing endonuclease I-SceI." J Am Chem Soc **128**(7): 2477-84.
Doyon, Y., J. M. McCammon, et al. (2008). "Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases
nbt1409 [pii] 10.1038/nbt1409." Nature Biotechnology **26**(6): 702-8.
Eastberg, J. H., J. Eklund, et al. (2007). "Mutability of an HNH nuclease imidazole general base and exchange of a deprotonation mechanism." Biochemistry **46**(24): 7215-25.
Eisenschmidt, K., T. Lanio, et al. (2005). "Developing a programmed restriction endonuclease for highly specific DNA cleavage." Nucleic Acids Res **33**(22): 7039-47.
Elrod-Erickson, M., M. A. Rould, et al. (1996). "Zif268 protein-DNA complex refined at 1.6 A: a model system for understanding zinc finger-DNA interactions." Structure **4**(10): 1171-80.
Epinat, J. C., S. Arnould, et al. (2003). "A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells." Nucleic Acids Research **31**(11): 2952-62.
Epinat, J. C., S. Arnould, et al. (2003). "A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells." Nucleic Acids Res **31**(11): 2952-62.
Frank, K. M., J. M. Sekiguchi, et al. (1998). "Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV
10.1038/24172." Nature **396**(6707): 173-7.
Galetto, R., P. Duchateau, et al. (2009). "Targeted approaches for gene therapy and the emergence of engineered meganucleases
10.1517/14712590903213669." Expert opinion on biological therapy **9**(10): 1289-303.
Gao, H., J. Smith, et al. (2010). "Heritable targeted mutagenesis in maize using a designed endonuclease
TPJ4041 [pii] 10.1111/j.1365-313X.2009.04041.x." The Plant journal : for cell and molecular biology **61**(1): 176-87.
Gao, Y., Y. Sun, et al. (1998). "A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis
S0092-8674(00)81714-6 [pii]." Cell **95**(7): 891-902.
Geurts, A. M., G. J. Cost, et al. (2009). "Knockout rats via embryo microinjection of zinc-finger nucleases
325/5939/433 [pii] 10.1126/science.1172447." Science **325**(5939): 433.
Gimble, F. S., C. M. Moure, et al. (2003). "Assessing the plasticity of DNA target site recognition of the PI-SceI homing endonuclease using a bacterial two-hybrid selection system." J Mol Biol **334**(5): 993-1008.
Greisman, H. A. and C. O. Pabo (1997). "A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites." Science **275**(5300): 657-61.
Gruenert, D. C., E. Bruscia, et al. (2003). "Sequence-specific modification of genomic DNA by small DNA fragments
10.1172/JCI19773 112/5/637 [pii]." The Journal of clinical investigation **112**(5): 637-41.
Guirouilh-Barbat, J., S. Huck, et al. (2004). "Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells." Mol Cell **14**(5): 611-23.
Guirouilh-Barbat, J., S. Huck, et al. (2004). "Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells
10.1016/j.molcel.2004.05.008 S1097276504002916 [pii]." Molecular Cell **14**(5): 611-23.
Guirouilh-Barbat, J., E. Rass, et al. (2007). "Defects in XRCC4 and KU80 differentially affect the joining of distal nonhomologous ends." Proc Natl Acad Sci U S A **104**(52): 20902-7.
Guirouilh-Barbat, J., E. Rass, et al. (2007). "Defects in XRCC4 and KU80 differentially affect the joining of distal nonhomologous ends

0708541104 [pii] 10.1073/pnas.0708541104." Proceedings of the National Academy of Sciences of the United States of America **104**(52): 20902-7.

Gurlebeck, D., B. Szurek, et al. (2005). "Dimerization of the bacterial effector protein AvrBs3 in the plant cell cytoplasm prior to nuclear import

TPJ2370 [pii] 10.1111/j.1365-313X.2005.02370.x." The Plant journal : for cell and molecular biology **42**(2): 175-87.

Haber, J. (2000). "Partners and pathwaysrepairing a double-strand break." Trends Genet. **16**(6): 259-264.

Haber, J. E. (2008). "Alternative endings

0711334105 [pii] 10.1073/pnas.0711334105." Proceedings of the National Academy of Sciences of the United States of America **105**(2): 405-6.

Hartsuiker, E., K. Mizuno, et al. (2009). "Ctp1CtIP and Rad32Mre11 nuclease activity are required for Rec12Spo11 removal, but Rec12Spo11 removal is dispensable for other MRN-dependent meiotic functions." Mol Cell Biol **29**(7): 1671-81.

Hinnen, A., J. B. Hicks, et al. (1978). "Transformation of yeast." Proceedings of the National Academy of Sciences of the United States of America **75**(4): 1929-33.

Hirata, R., J. Chamberlain, et al. (2002). "Targeted transgene insertion into human chromosomes by adeno-associated virus vectors

10.1038/nbt0702-735 nbt0702-735 [pii]." Nature Biotechnology **20**(7): 735-8.

Huang, H. and H. S. Yuan (2007). "The conserved asparagine in the HNH motif serves an important structural role in metal finger endonucleases." Journal of Molecular Biology **368**(3): 812-21.

Ichiyanagi, K., Y. Ishino, et al. (2000). "Crystal structure of an archaeal intein-encoded homing endonuclease PI-Pful." J Mol Biol **300**(4): 889-901.

Inoue, N., R. Dong, et al. (2001). "Introduction of single base substitutions at homologous chromosomal sequences by adeno-associated virus vectors

10.1006/mthe.2001.0283 S1525-0016(01)90283-7 [pii]." Molecular therapy : the journal of the American Society of Gene Therapy **3**(4): 526-30.

Isalan, M. and Y. Choo (2001). "Rapid, high-throughput engineering of sequence-specific zinc finger DNA-binding proteins

S0076-6879(01)40444-7 [pii]." Methods in Enzymology **340**: 593-609.

Kalish, J. M. and P. M. Glazer (2005). "Targeted genome modification via triple helix formation." Ann N Y Acad Sci **1058**: 151-61.

Kim, H. J., H. J. Lee, et al. (2009). "Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly

gr.089417.108 [pii] 10.1101/gr.089417.108." Genome Research **19**(7): 1279-88.

Kim, Y. G., J. Cha, et al. (1996). "Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain." Proceedings of the National Academy of Sciences of the United States of America **93**(3): 1156-60.

Ku, W. Y., Y. W. Liu, et al. (2002). "The zinc ion in the HNH motif of the endonuclease domain of colicin E7 is not required for DNA binding but is essential for DNA hydrolysis." Nucleic Acids Research **30**(7): 1670-8.

Landthaler, M., U. Begley, et al. (2002). "Two self-splicing group I introns in the ribonucleotide reductase large subunit gene of Staphylococcus aureus phage Twort." Nucleic Acids Research **30**(9): 1935-43.

Landthaler, M., N. C. Lau, et al. (2004). "Group I intron homing in Bacillus phages SPO1 and SP82: a gene conversion event initiated by a nicking homing endonuclease." Journal of Bacteriology **186**(13): 4307-14.

Landthaler, M., B. W. Shen, et al. (2006). "I-BasI and I-HmuI: two phage intron-encoded endonucleases with homologous DNA recognition sequences but distinct DNA specificities." Journal of Molecular Biology **358**(4): 1137-51.

Landthaler, M. and D. A. Shub (2003). "The nicking homing endonuclease I-BasI is encoded by a group I intron in the DNA polymerase gene of the Bacillus thuringiensis phage Bastille." Nucleic Acids Research **31**(12): 3071-7.

Lee, S. E., F. Paques, et al. (1999). "Role of yeast SIR genes and mating type in directing DNA double-strand breaks to homologous and non-homologous repair paths." Curr Biol **9**(14): 767-70.

Li, T., S. Huang, et al. (2010). "TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain." Nucleic Acids Res **39**(1): 359-72.

Li, T., S. Huang, et al. (2011). "TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain
gkq704 [pii] 10.1093/nar/gkq704." Nucleic Acids Research **39**(1): 359-72.

Liang, F., M. Han, et al. (1998). "Homology-directed repair is a major double-strand break repair pathway in mammalian cells." Proceedings of the National Academy of Sciences of the United States of America **95**(9): 5172-7.

Liu, P. Q., E. M. Chan, et al. (2010). "Generation of a triple-gene knockout mammalian cell line using engineered zinc-finger nucleases
10.1002/bit.22654." Biotechnology and Bioengineering **106**(1): 97-105.

Liu, Q., J. T. Dansereau, et al. (2008). "Role of the interdomain linker in distance determination for remote cleavage by homing endonuclease I-TevI." J Mol Biol **379**(5): 1094-106.

Liu, Q., V. Derbyshire, et al. (2006). "Distance determination by GIY-YIG intron endonucleases: discrimination between repression and cleavage functions." Nucleic Acids Research **34**(6): 1755-64.

Lloyd, A., C. L. Plaisier, et al. (2005). "Targeted mutagenesis using zinc-finger nucleases in Arabidopsis
0409339102 [pii] 10.1073/pnas.0409339102." Proceedings of the National Academy of Sciences of the United States of America **102**(6): 2232-7.

Maeder, M. L., S. Thibodeau-Beganny, et al. (2008). "Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification
S1097-2765(08)00461-9 [pii] 10.1016/j.molcel.2008.06.016." Molecular Cell **31**(2): 294-301.

Mahfouz, M. M., L. Li, et al. (2011). "De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks
1019533108 [pii] 10.1073/pnas.1019533108." Proceedings of the National Academy of Sciences of the United States of America **108**(6): 2623-8.

Marcaida, M. J., I. G. Munoz, et al. (2010). "Homing endonucleases: from basics to therapeutic applications
10.1007/s00018-009-0188-y." Cellular and molecular life sciences : CMLS **67**(5): 727-48.

Mashimo, T., A. Takizawa, et al. (2010). "Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases
10.1371/journal.pone.0008870." PloS one **5**(1): e8870.

McConnell Smith, A., R. Takeuchi, et al. (2009). "Generation of a nicking enzyme that stimulates site-specific gene conversion from the I-AniI LAGLIDADG homing endonuclease
0810588106 [pii] 10.1073/pnas.0810588106." Proceedings of the National Academy of Sciences of the United States of America **106**(13): 5099-104.

McVey, M. and S. E. Lee (2008). "MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings
S0168-9525(08)00229-1 [pii] 10.1016/j.tig.2008.08.007." Trends in genetics : TIG **24**(11): 529-38.

Menoret, S., A. L. Iscache, et al. (2010). "Characterization of immunoglobulin heavy chain knockout rats
10.1002/eji.201040939." European Journal of Immunology **40**(10): 2932-41.

Metzger, M. J., A. McConnell-Smith, et al. (2011). "Single-strand nicks induce homologous recombination with less toxicity than double-strand breaks using an AAV vector template
gkq826 [pii] 10.1093/nar/gkq826." Nucleic Acids Research **39**(3): 926-35.

Midon, M., P. Schafer, et al. (2011). "Mutational and biochemical analysis of the DNA-entry nuclease EndA from Streptococcus pneumoniae
gkq802 [pii] 10.1093/nar/gkq802." Nucleic Acids Research **39**(2): 623-34.
Miller, J. C., S. Tan, et al. (2011). "A TALE nuclease architecture for efficient genome editing
nbt.1755 [pii] 10.1038/nbt.1755." Nature Biotechnology **29**(2): 143-8.
Mimitou, E. P. and L. S. Symington (2008). "Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing." Nature **455**(7214): 770-4.
Moore, I., M. Samalova, et al. (2006). "Transactivated and chemically inducible gene expression in plants." Plant J **45**(4): 651-83.
Moore, J. K. and J. E. Haber (1996). "Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in Saccharomyces cerevisiae." Mol Cell Biol **16**(5): 2164-73.
Moscou, M. J. and A. J. Bogdanove (2009). "A simple cipher governs DNA recognition by TAL effectors." Science **326**(5959): 1501.
Moscou, M. J. and A. J. Bogdanove (2009). "A simple cipher governs DNA recognition by TAL effectors
1178817 [pii] 10.1126/science.1178817." Science **326**(5959): 1501.
Moure, C. M., F. S. Gimble, et al. (2002). "Crystal structure of the intein homing endonuclease PI-SceI bound to its recognition sequence." Nat Struct Biol **9**(10): 764-70.
Moure, C. M., F. S. Gimble, et al. (2003). "The crystal structure of the gene targeting homing endonuclease I-SceI reveals the origins of its target site specificity." J Mol Biol **334**(4): 685-95.
Nimonkar, A. V., J. Genschel, et al. (2011). "BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair." Genes Dev **25**(4): 350-62.
Niu, Y., K. Tenney, et al. (2008). "Engineering variants of the I-SceI homing endonuclease with strand-specific and site-specific DNA-nicking activity
S0022-2836(08)00840-1 [pii] 10.1016/j.jmb.2008.07.010." Journal of Molecular Biology **382**(1): 188-202.
Orr-Weaver, T. L., J. W. Szostak, et al. (1981). "Yeast transformation: a model system for the study of recombination." Proceedings of the National Academy of Sciences of the United States of America **78**(10): 6354-8.
Orr-Weaver, T. L., J. W. Szostak, et al. (1983). "Genetic applications of yeast transformation with linear and gapped plasmids." Methods in Enzymology **101**: 228-45.
Pabo, C. O., E. Peisach, et al. (2001). "Design and selection of novel Cys2His2 zinc finger proteins
70/1/313 [pii] 10.1146/annurev.biochem.70.1.313." Annual Review of Biochemistry **70**: 313-40.
Padidam, M. (2003). "Chemically regulated gene expression in plants." Curr Opin Plant Biol **6**(2): 169-77.
Paques, F. and P. Duchateau (2007). "Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy." Current Gene Therapy **7**(1): 49-66.
Paques, F. and P. Duchateau (2007). "Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy." Curr Gene Ther **7**(1): 49-66.
Paques, F. and J. E. Haber (1999). "Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae." Microbiology and molecular biology reviews : MMBR **63**(2): 349-404.
Perez, E. E., J. Wang, et al. (2008). "Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases
nbt1410 [pii] 10.1038/nbt1410." Nature Biotechnology **26**(7): 808-16.
Pierce, A. J., P. Hu, et al. (2001). "Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells." Genes Dev **15**(24): 3237-42.
Pingoud, A. and G. H. Silva (2007). "Precision genome surgery." Nat Biotechnol **25**(7): 743-4.

Porteus, M. H. and D. Carroll (2005). "Gene targeting using zinc finger nucleases." Nat Biotechnol **23**(8): 967-73.

Ramirez, C. L., J. E. Foley, et al. (2008). "Unexpected failure rates for modular assembly of engineered zinc fingers
nmeth0508-374 [pii] 10.1038/nmeth0508-374." Nature Methods **5**(5): 374-5.

Rosen, L. E., H. A. Morrison, et al. (2006). "Homing endonuclease I-CreI derivatives with novel DNA target specificities." Nucleic Acids Research **34**(17): 4791-800.

Rosen, L. E., H. A. Morrison, et al. (2006). "Homing endonuclease I-CreI derivatives with novel DNA target specificities." Nucleic Acids Res.

Rothstein, R. J. (1983). "One-step gene disruption in yeast." Methods in Enzymology **101**: 202-11.

Rouet, P., F. Smih, et al. (1994). "Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells." Proc Natl Acad Sci U S A **91**(13): 6064-8.

Rouet, P., F. Smih, et al. (1994). "Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease." Mol Cell Biol **14**(12): 8096-106.

Russell, D. W. and R. K. Hirata (1998). "Human gene targeting by viral vectors
10.1038/ng0498-325." Nature Genetics **18**(4): 325-30.

Sangiuolo, F., M. L. Scaldaferri, et al. (2008). "Cftr gene targeting in mouse embryonic stem cells mediated by Small Fragment Homologous Replacement (SFHR)
2904 [pii]." Frontiers in bioscience : a journal and virtual library **13**: 2989-99.

Santiago, Y., E. Chan, et al. (2008). "Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases
0800940105 [pii] 10.1073/pnas.0800940105." Proceedings of the National Academy of Sciences of the United States of America **105**(15): 5809-14.

Sartori, A. A., C. Lukas, et al. (2007). "Human CtIP promotes DNA end resection." Nature **450**(7169): 509-14.

Seligman, L. M., K. M. Chisholm, et al. (2002). "Mutations altering the cleavage specificity of a homing endonuclease." Nucleic Acids Research **30**(17): 3870-9.

Seligman, L. M., K. M. Stephens, et al. (1997). "Genetic analysis of the Chlamydomonas reinhardtii I-CreI mobile intron homing system in Escherichia coli." Genetics **147**(4): 1653-64.

Shen, B. W., M. Landthaler, et al. (2004). "DNA binding and cleavage by the HNH homing endonuclease I-HmuI." Journal of Molecular Biology **342**(1): 43-56.

Shukla, V. K., Y. Doyon, et al. (2009). "Precise genome modification in the crop species Zea mays using zinc-finger nucleases
nature07992 [pii] 10.1038/nature07992." Nature **459**(7245): 437-41.

Silva, G. H., J. Z. Dalgaard, et al. (1999). "Crystal structure of the thermostable archaeal intron-encoded endonuclease I-DmoI." J Mol Biol **286**(4): 1123-36.

Simon, P., F. Cannata, et al. (2008). "Sequence-specific DNA cleavage mediated by bipyridine polyamide conjugates." Nucleic Acids Res **36**(11): 3531-8.

Smith, J., J. M. Berg, et al. (1999). "A detailed study of the substrate specificity of a chimeric restriction enzyme
gkc139 [pii]." Nucleic Acids Research **27**(2): 674-81.

Smith, J., M. Bibikova, et al. (2000). "Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains." Nucleic Acids Research **28**(17): 3361-9.

Smith, J., S. Grizot, et al. (2006). "A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences." Nucleic Acids Research **34**(22): e149.

Smith, J., S. Grizot, et al. (2006). "A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences." Nucleic Acids Res **34**(22): e149.

Sonoda, E., H. Hochegger, et al. (2006). "Differential usage of non-homologous end-joining and homologous recombination in double strand break repair." DNA Repair (Amst) **5**(9-10): 1021-9.

Spiegel, P. C., B. Chevalier, et al. (2006). "The structure of I-CeuI homing endonuclease: Evolving asymmetric DNA recognition from a symmetric protein scaffold." Structure **14**(5): 869-80.

Stoddard, B. L. (2005). "Homing endonuclease structure and function." Quarterly Reviews of Biophysics **38**(1): 49-95.

Stoddard, B. L. (2005). "Homing endonuclease structure and function." Q Rev Biophys **38**(1): 49-95.

Stoddard, B. L., A. M. Scharenberg, et al. (2007). Advances in Engineering Homing Endonucleases for Gene Targeting: Ten Years After Structures. Progress in Gene Therapy: Autologous and Cancer Stem Cell Gene Therapy. R. Bertolotti and K. Ozawa, World Scientific Publishing Co. Pte. Ltd. **3:** 135-68.

Sugawara, N. and J. E. Haber (1992). "Characterization of double-strand break-induced recombination: homology requirements and single-stranded DNA formation." Mol Cell Biol **12**(2): 563-75.

Sun, H., D. Treco, et al. (1991). "Extensive 3'-overhanging, single-stranded DNA associated with the meiosis-specific double-strand breaks at the ARG4 recombination initiation site." Cell **64**(6): 1155-61.

Sussman, D., M. Chadsey, et al. (2004). "Isolation and characterization of new homing endonuclease specificities at individual target site positions." Journal of Molecular Biology **342**(1): 31-41.

Sussman, D., M. Chadsey, et al. (2004). "Isolation and characterization of new homing endonuclease specificities at individual target site positions." J Mol Biol **342**(1): 31-41.

Taubes, G. (2002). "Gene therapy. The strange case of chimeraplasty
10.1126/science.298.5601.2116 298/5601/2116 [pii]." Science **298**(5601): 2116-20.

Wang, R., X. Zhou, et al. (2003). "Chemically regulated expression systems and their applications in transgenic plants." Transgenic Res **12**(5): 529-40.

Wang, Y. T., J. D. Wright, et al. (2009). "Redesign of high-affinity nonspecific nucleases with altered sequence preference
10.1021/ja907160r." Journal of the American Chemical Society **131**(47): 17345-53.

White, C. I. and J. E. Haber (1990). "Intermediates of recombination during mating type switching in Saccharomyces cerevisiae." Embo J **9**(3): 663-73.

Yang, M., V. Djukanovic, et al. (2009). "Targeted mutagenesis in the progeny of maize transgenic plants
10.1007/s11103-009-9499-5." Plant Molecular Biology **70**(6): 669-79.

Zhao, L., R. P. Bonocora, et al. (2007). "The restriction fold turns to the dark side: a bacterial homing endonuclease with a PD-(D/E)-XK motif." The EMBO Journal **26**(9): 2432-42.

Zuo, J. and N. H. Chua (2000). "Chemical-inducible systems for regulated expression of plant genes." Curr Opin Biotechnol **11**(2): 146-51.

Claims

1) A compact TALEN monomer comprising:

(i) One core TALE scaffold comprising Repeat Variable Dipeptide regions (RVDs) having DNA binding specificity onto a specific double-stranded DNA target sequence of interest;

(ii) At least one catalytic domain of I-TevI capable of cleaving DNA adjacent to said double-stranded DNA target sequence of interest when fused to the C or N terminal of said core TALE scaffold from (i);

wherein said compact TALEN monomer binds said target DNA sequence and cleaves double-stranded DNA.

2) A compact TALEN monomer of claim 1 further comprising one peptidic linker to fuse said catalytic domain from (ii) to said core TALE scaffold from (i).

3) A compact TALEN monomer according to claim 1 or 2, wherein said catalytic domain is fused to the C-terminal domain of said core TALE scaffold.

4) A compact TALEN monomer according to claim 1 or 2, wherein said catalytic domain is fused to the N-terminal domain of said core TALE scaffold.

5) A compact TALEN monomer according to claims 1 to 4, comprising a protein sequence having at least 80%, more preferably 90%, again more preferably 95% amino acid sequence identity with the protein sequences selected from the group of SEQ ID NO: 426-432.

6) A compact TALEN monomer according to any one of claims 1 to 5, wherein said peptidic linker sequence can be selected from the group consisting of SEQ ID NO: 67-104 and SEQ ID NO: 372 to SEQ ID NO: 415.

7) A recombinant polynucleotide encoding a compact TALEN according to any one of claims 1 to 6.

8) A vector comprising a recombinant polynucleotide according to claim 7.

9) A pharmaceutical composition comprising a compact TALEN according to any one of claim 1 to 6 and a pharmaceutically active carrier.

10)A host cell which comprises a recombinant polynucleotide of claim 7.

11) A non-human transgenic animal which comprises a recombinant polynucleotide of claim 7.

12) A transgenic plant which comprises a recombinant polynucleotide of claim 7.

13) A kit comprising a compact TALEN monomer according to any one of claims 1 to 6 and instructions for use in enhancing DNA processing efficiency of a single double-stranded DNA target sequence of interest.

14) A method for targeting and cleaving a double-stranded DNA, comprising:

(a) Selecting one DNA target sequence of interest on one strand of a double-stranded DNA;

(b) Providing a compact TALEN monomer according to any one of claims 1 to 6;

(c) Contacting said double-stranded DNA with said monomer such that said compact TALEN monomer cleaves said double stranded DNA.

15) A method for inserting a transgene into a specific single double-stranded DNA target sequence of a genomic locus of a cell, tissue or non-human animal wherein at least one compact TALEN monomer of any one of claims 1 to 6 is introduced in said cell, tissue or non-human animal.

SEQUENCE LISTING

<110> SILVA GEORGE
BERTONATI CLAUDIA
EPINAT JEAN-CHARLES
DUCHATEAU PHILIPPE
VALTON JULIEN
JUILLERAT ALEXANDRE

<120> METHOD FOR THE GENERATION OF COMPACT TALE-NUCLEASES AND USES THEREOF

<130> 397357WO

<160> 455

<170> PatentIn version 3.5

<210> 1
<211> 163
<212> PRT
<213> Chlamydomonas reinhardtii

<400> 1

Met Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe
1 5 10 15

Val Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser
20 25 30

Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys
35 40 45

Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val
50 55 60

Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu
65 70 75 80

Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys
85 90 95

Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu
100 105 110

Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp
115 120 125

Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr
130 135 140

Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys
145 150 155 160

Ser Ser Pro

<210> 2
<211> 22
<212> DNA
<213> Artificial

<220>
<223> C1221 target

<400> 2
caaaacgtcg tacgacgttt tg 22

<210> 3
<211> 22
<212> DNA
<213> Artificial

<220>
<223> C1234 target

<400> 3
caaaacgtcg tgagacagtt tg 22

<210> 4

<211> 412
<212> PRT
<213> artificial sequence

<220>
<223> fem-3 gi|262368139|pdb|3K5Q|

<400> 4
Ser Asn Asn Val Leu Pro Thr Trp Ser Leu Asp Ser Asn Gly Glu Met
1 5 10 15

Arg Ser Arg Leu Ser Leu Ser Glu Val Leu Asp Ser Gly Asp Leu Met
20 25 30

Lys Phe Ala Val Asp Lys Thr Gly Cys Gln Phe Leu Glu Lys Ala Val
35 40 45

Lys Gly Ser Leu Thr Ser Tyr Gln Lys Phe Gln Leu Phe Glu Gln Val
50 55 60

Ile Gly Arg Lys Asp Asp Phe Leu Lys Leu Ser Thr Asn Ile Phe Gly
65 70 75 80

Asn Tyr Leu Val Gln Ser Val Ile Gly Ile Ser Leu Ala Thr Asn Asp
85 90 95

Asp Gly Tyr Thr Lys Arg Gln Glu Lys Leu Lys Asn Phe Ile Ser Ser
100 105 110

Gln Met Thr Asp Met Cys Leu Asp Lys Phe Ala Cys Arg Val Ile Gln
115 120 125

Ser Ser Leu Gln Asn Met Asp Leu Ser Leu Ala Cys Lys Leu Val Gln
130 135 140

```
Ala Leu Pro Arg Asp Ala Arg Leu Ile Ala Ile Cys Val Asp Gln Asn
145                 150                 155                 160

Ala Asn His Val Ile Gln Lys Val Val Ala Val Ile Pro Leu Lys Asn
                165                 170                 175

Trp Glu Phe Ile Val Asp Phe Val Ala Thr Pro Glu His Leu Arg Gln
            180                 185                 190

Ile Cys Ser Asp Lys Tyr Gly Cys Arg Val Val Gln Thr Ile Ile Glu
        195                 200                 205

Lys Leu Thr Ala Asp Ser Met Asn Val Asp Leu Thr Ser Ala Ala Gln
    210                 215                 220

Asn Leu Arg Glu Arg Ala Leu Gln Arg Leu Met Thr Ser Val Thr Asn
225                 230                 235                 240

Arg Cys Gln Glu Leu Ala Thr Asn Glu Tyr Ala Asn Tyr Ile Ile Gln
                245                 250                 255

His Ile Val Ser Asn Asp Asp Leu Ala Val Tyr Arg Glu Cys Ile Ile
            260                 265                 270

Glu Lys Cys Leu Met Arg Asn Leu Leu Ser Leu Ser Gln Glu Lys Phe
        275                 280                 285

Ala Ser His Val Val Glu Lys Ala Phe Leu His Ala Pro Leu Glu Leu
    290                 295                 300

Leu Ala Glu Met Met Asp Glu Ile Phe Asp Gly Tyr Ile Pro His Pro
305                 310                 315                 320

Asp Thr Gly Lys Asp Ala Leu Asp Ile Met Met Phe His Gln Phe Gly
                325                 330                 335

Asn Tyr Val Val Gln Cys Met Leu Thr Ile Cys Cys Asp Ala Val Ser
            340                 345                 350

Gly Arg Arg Gln Thr Lys Glu Gly Gly Tyr Asp His Ala Ile Ser Phe
        355                 360                 365

Gln Asp Trp Leu Lys Lys Leu His Ser Arg Val Thr Lys Glu Arg His
    370                 375                 380

Arg Leu Ser Arg Phe Ser Ser Gly Lys Lys Met Ile Glu Thr Leu Ala
385                 390                 395                 400

Asn Leu Arg Ser Thr His Pro Ile Tyr Glu Leu Gln
                405                 410
```

<210> 5

<211> 201
<212> PRT
<213> artificial sequence

<220>
<223> aRep gi|308387836|pdb|3LTJ|

<400> 5
Met Arg Gly Ser His His His His His His Thr Asp Pro Glu Lys Val
1 5 10 15

Glu Met Tyr Ile Lys Asn Leu Gln Asp Asp Ser Tyr Tyr Val Arg Arg
20 25 30

Ala Ala Ala Tyr Ala Leu Gly Lys Ile Gly Asp Glu Arg Ala Val Glu
35 40 45

Pro Leu Ile Lys Ala Leu Lys Asp Glu Asp Ala Trp Val Arg Arg Ala
50 55 60

Ala Ala Asp Ala Leu Gly Gln Ile Gly Asp Glu Arg Ala Val Glu Pro
65 70 75 80

Leu Ile Lys Ala Leu Lys Asp Glu Asp Gly Trp Val Arg Gln Ser Ala
85 90 95

Ala Val Ala Leu Gly Gln Ile Gly Asp Glu Arg Ala Val Glu Pro Leu
100 105 110

Ile Lys Ala Leu Lys Asp Glu Asp Trp Phe Val Arg Ile Ala Ala Ala
115 120 125

Phe Ala Leu Gly Glu Ile Gly Asp Glu Arg Ala Val Glu Pro Leu Ile
130 135 140

Lys Ala Leu Lys Asp Glu Asp Gly Trp Val Arg Gln Ser Ala Ala Asp
145 150 155 160

Ala Leu Gly Glu Ile Gly Gly Glu Arg Val Arg Ala Ala Met Glu Lys
165 170 175

Leu Ala Glu Thr Gly Thr Gly Phe Ala Arg Lys Val Ala Val Asn Tyr
180 185 190

Leu Glu Thr His Lys Ser Leu Ile Ser
195 200

<210> 6

<211> 309
<212> PRT

<213> Pseudomonas Aeruginosa

<220>
<223> Ligd Polymerase Domain gi|109157579|pdb|2FAO|

<400> 6

```
Met Gly Ala Arg Lys Ala Ser Ala Gly Ala Ser Arg Ala Ala Thr Ala
1               5                  10                  15

Gly Val Arg Ile Ser His Pro Gln Arg Leu Ile Asp Pro Ser Ile Gln
            20                  25                  30

Ala Ser Lys Leu Glu Leu Ala Glu Phe His Ala Arg Tyr Ala Asp Leu
        35                  40                  45

Leu Leu Arg Asp Leu Arg Glu Arg Pro Val Ser Leu Val Arg Gly Pro
    50                  55                  60

Asp Gly Ile Gly Gly Glu Leu Phe Phe Gln Lys His Ala Ala Arg Leu
65                  70                  75                  80

Lys Ile Pro Gly Ile Val Gln Leu Asp Pro Ala Leu Asp Pro Gly His
                85                  90                  95

Pro Pro Leu Leu Gln Ile Arg Ser Ala Glu Ala Leu Val Gly Ala Val
            100                 105                 110

Gln Met Gly Ser Ile Glu Phe His Thr Trp Asn Ala Ser Leu Ala Asn
        115                 120                 125

Leu Glu Arg Pro Asp Arg Phe Val Leu Asp Leu Asp Pro Asp Pro Ala
    130                 135                 140

Leu Pro Trp Lys Arg Met Leu Glu Ala Thr Gln Leu Ser Leu Thr Leu
145                 150                 155                 160

Leu Asp Glu Leu Gly Leu Arg Ala Phe Leu Lys Thr Ser Gly Gly Lys
                165                 170                 175

Gly Met His Leu Leu Val Pro Leu Glu Arg Arg His Gly Trp Asp Glu
            180                 185                 190

Val Lys Asp Phe Ala Gln Ala Ile Ser Gln His Leu Ala Arg Leu Met
        195                 200                 205

Pro Glu Arg Phe Ser Ala Val Ser Gly Pro Arg Asn Arg Val Gly Lys
    210                 215                 220

Ile Phe Val Asp Tyr Leu Arg Asn Ser Arg Gly Ala Ser Thr Val Ala
225                 230                 235                 240

Ala Tyr Ser Val Arg Ala Arg Glu Gly Leu Pro Val Ser Val Pro Val
                245                 250                 255
```

Phe Arg Glu Glu Leu Asp Ser Leu Gln Gly Ala Asn Gln Trp Asn Leu
260 265 270

Arg Ser Leu Pro Gln Arg Leu Asp Glu Leu Ala Gly Asp Asp Pro Trp
275 280 285

Ala Asp Tyr Ala Gly Thr Arg Gln Arg Ile Ser Ala Ala Met Arg Arg
290 295 300

Gln Leu Gly Arg Gly
305

<210> 7

<211> 303
<212> PRT
<213> Mycobacterium Tuberculosis

<220>
<223> Ligase D 2R9L_A GI: 164519498

<400> 7
Gly Ser His Met Gly Ser Ala Ser Glu Gln Arg Val Thr Leu Thr Asn
1 5 10 15

Ala Asp Lys Val Leu Tyr Pro Ala Thr Gly Thr Thr Lys Ser Asp Ile
20 25 30

Phe Asp Tyr Tyr Ala Gly Val Ala Glu Val Met Leu Gly His Ile Ala
35 40 45

Gly Arg Pro Ala Thr Arg Lys Arg Trp Pro Asn Gly Val Asp Gln Pro
50 55 60

Ala Phe Phe Glu Lys Gln Leu Ala Leu Ser Ala Pro Pro Trp Leu Ser
65 70 75 80

Arg Ala Thr Val Ala His Arg Ser Gly Thr Thr Thr Tyr Pro Ile Ile
85 90 95

Asp Ser Ala Thr Gly Leu Ala Trp Ile Ala Gln Gln Ala Ala Leu Glu
100 105 110

Val His Val Pro Gln Trp Arg Phe Val Ala Glu Pro Gly Ser Gly Glu
115 120 125

Leu Asn Pro Gly Pro Ala Thr Arg Leu Val Phe Asp Leu Asp Pro Gly
130 135 140

Glu Gly Val Met Met Ala Gln Leu Ala Glu Val Ala Arg Ala Val Arg
145 150 155 160

```
Asp Leu Leu Ala Asp Ile Gly Leu Val Thr Phe Pro Val Thr Ser Gly
                165                 170                 175

Ser Lys Gly Leu His Leu Tyr Thr Pro Leu Asp Glu Pro Val Ser Ser
            180                 185                 190

Arg Gly Ala Thr Val Leu Ala Lys Arg Val Ala Gln Arg Leu Glu Gln
        195                 200                 205

Ala Met Pro Ala Leu Val Thr Ser Thr Met Thr Lys Ser Leu Arg Ala
    210                 215                 220

Gly Lys Val Phe Val Asp Trp Ser Gln Asn Ser Gly Ser Lys Thr Thr
225                 230                 235                 240

Ile Ala Pro Tyr Ser Leu Arg Gly Arg Thr His Pro Thr Val Ala Ala
                245                 250                 255

Pro Arg Thr Trp Ala Glu Leu Asp Asp Pro Ala Leu Arg Gln Leu Ser
            260                 265                 270

Tyr Asp Glu Val Leu Thr Arg Ile Ala Arg Asp Gly Asp Leu Leu Glu
        275                 280                 285

Arg Leu Asp Ala Asp Ala Pro Val Ala Asp Arg Leu Thr Arg Tyr
    290                 295                 300
```

<210> 8

<211> 410
<212> PRT
<213> Pyrococcus Abyssi

<220>
<223> Large Gamma Subunit Of Initiation Factor Eif2 1KJZ_A GI:20664108

<400> 8

```
Gly Glu Lys Arg Lys Ser Arg Gln Ala Glu Val Asn Ile Gly Met Val
1               5                   10                  15

Gly His Val Asp His Gly Lys Thr Thr Leu Thr Lys Ala Leu Thr Gly
            20                  25                  30

Val Trp Thr Asp Thr His Ser Glu Glu Leu Arg Arg Gly Ile Thr Ile
        35                  40                  45

Lys Ile Gly Phe Ala Asp Ala Glu Ile Arg Arg Cys Pro Asn Cys Gly
    50                  55                  60

Arg Tyr Ser Thr Ser Pro Val Cys Pro Tyr Cys Gly His Glu Thr Glu
65                  70                  75                  80

Phe Val Arg Arg Val Ser Phe Ile Asp Ala Pro Gly His Glu Ala Leu
                85                  90                  95
```

```
Met Thr Thr Met Leu Ala Gly Ala Ser Leu Met Asp Gly Ala Ile Leu
            100                 105                 110

Val Ile Ala Ala Asn Glu Pro Cys Pro Arg Pro Gln Thr Arg Glu His
        115                 120                 125

Leu Met Ala Leu Gln Ile Ile Gly Gln Lys Asn Ile Ile Ile Ala Gln
    130                 135                 140

Asn Lys Ile Glu Leu Val Asp Lys Glu Lys Ala Leu Glu Asn Tyr Arg
145                 150                 155                 160

Gln Ile Lys Glu Phe Ile Glu Gly Thr Val Ala Glu Asn Ala Pro Ile
                165                 170                 175

Ile Pro Ile Ser Ala Leu His Gly Ala Asn Ile Asp Val Leu Val Lys
            180                 185                 190

Ala Ile Glu Asp Phe Ile Pro Thr Pro Lys Arg Asp Pro Asn Lys Pro
        195                 200                 205

Pro Lys Met Leu Val Leu Arg Ser Phe Asp Val Asn Lys Pro Gly Thr
    210                 215                 220

Pro Pro Glu Lys Leu Val Gly Gly Val Leu Asp Gly Ser Ile Val Gln
225                 230                 235                 240

Gly Lys Leu Lys Val Gly Asp Glu Ile Glu Ile Arg Pro Gly Val Pro
                245                 250                 255

Tyr Glu Glu His Gly Arg Ile Lys Tyr Glu Pro Ile Thr Thr Glu Ile
            260                 265                 270

Val Ser Leu Gln Ala Gly Gly Gln Phe Val Glu Glu Ala Tyr Pro Gly
        275                 280                 285

Gly Leu Val Gly Val Gly Thr Lys Leu Asp Pro Tyr Leu Thr Lys Gly
    290                 295                 300

Asp Leu Met Ala Gly Asn Val Val Gly Lys Pro Gly Lys Leu Pro Pro
305                 310                 315                 320

Val Trp Asp Ser Leu Arg Leu Glu Val His Leu Leu Glu Arg Val Val
                325                 330                 335

Gly Thr Glu Gln Glu Leu Lys Val Glu Pro Ile Lys Arg Lys Glu Val
            340                 345                 350

Leu Leu Leu Asn Val Gly Thr Ala Arg Thr Met Gly Leu Val Thr Gly
        355                 360                 365
```

Leu Gly Lys Asp Glu Ile Glu Val Lys Leu Gln Ile Pro Val Cys Ala
370 375 380

Glu Pro Gly Asp Arg Val Ala Ile Ser Arg Gln Ile Gly Ser Arg Trp
385 390 395 400

Arg Leu Ile Gly Tyr Gly Ile Ile Lys Glu
405 410

<210> 9

<211> 419
<212> PRT
<213> artificial sequence

<220>
<223> Translation Initiation Factor Aif2betagamma 2D74_A GI:112490420

<400> 9
Met Gly Glu Lys Arg Lys Thr Arg Gln Ala Glu Val Asn Ile Gly Met
1 5 10 15

Val Gly His Val Asp His Gly Lys Thr Thr Leu Thr Lys Ala Leu Thr
20 25 30

Gly Val Trp Thr Asp Thr His Ser Glu Glu Leu Arg Arg Gly Ile Thr
35 40 45

Ile Lys Ile Gly Phe Ala Asp Ala Glu Ile Arg Arg Cys Ser Asn Cys
50 55 60

Gly Arg Tyr Ser Thr Ser Pro Ile Cys Pro Tyr Cys Gly His Glu Thr
65 70 75 80

Glu Phe Ile Arg Arg Val Ser Phe Ile Asp Ser Pro Gly His Glu Ala
85 90 95

Leu Met Thr Thr Met Leu Ala Gly Ala Ser Leu Met Asp Gly Ala Ile
100 105 110

Leu Val Ile Ala Ala Asn Glu Pro Cys Pro Arg Pro Gln Thr Arg Glu
115 120 125

His Leu Met Ala Leu Gln Ile Ile Gly Gln Lys Asn Ile Ile Ile Ala
130 135 140

Gln Asn Lys Ile Glu Leu Val Asp Lys Glu Lys Ala Leu Glu Asn Tyr
145 150 155 160

Arg Gln Ile Lys Glu Phe Ile Lys Gly Thr Val Ala Glu Asn Ala Pro
165 170 175

Ile Ile Pro Ile Ser Ala Leu His Gly Ala Asn Ile Asp Val Leu Val
180 185 190

Lys Ala Ile Glu Glu Phe Ile Pro Thr Pro Lys Arg Asp Ser Asn Lys
195 200 205

Pro Pro Lys Met Leu Val Leu Arg Ser Phe Asp Val Asn Lys Pro Gly
210 215 220

Thr Pro Pro Glu Lys Leu Val Gly Gly Val Leu Asp Gly Ser Ile Val
225 230 235 240

Gln Gly Lys Leu Lys Val Gly Asp Glu Ile Glu Ile Arg Pro Gly Val
245 250 255

Pro Tyr Glu Glu His Gly Arg Ile Lys Tyr Glu Pro Ile Thr Thr Glu
260 265 270

Ile Val Ser Leu Gln Ala Gly Gly Gln Phe Val Glu Glu Ala Tyr Pro
275 280 285

Gly Gly Leu Val Gly Ile Gly Thr Lys Leu Asp Pro Tyr Leu Thr Lys
290 295 300

Gly Asp Leu Met Ala Gly Asn Val Val Gly Lys Pro Gly Lys Leu Pro
305 310 315 320

Pro Val Trp Thr Asp Leu Arg Leu Glu Val His Leu Leu Glu Arg Val
325 330 335

Val Gly Thr Glu Gln Glu Leu Asn Val Glu Pro Ile Lys Arg Lys Glu
340 345 350

Val Leu Leu Leu Asn Val Gly Thr Ala Arg Thr Met Gly Leu Val Thr
355 360 365

Ala Leu Gly Lys Asp Glu Ile Glu Leu Lys Leu Gln Ile Pro Val Cys
370 375 380

Ala Glu Pro Gly Glu Arg Val Ala Ile Ser Arg Gln Ile Gly Ser Arg
385 390 395 400

Trp Arg Leu Ile Gly Tyr Gly Ile Ile Lys Glu Leu Glu His His His
405 410 415

His His His

<210> 10

<211> 919
<212> PRT
<213> Methylophilus methylotrophus

<220>
<223> MmeI ACC85607.1

<400> 10
Met Ala Leu Ser Trp Asn Glu Ile Arg Arg Lys Ala Ile Glu Phe Ser
1   5   10   15

Lys Arg Trp Glu Asp Ala Ser Asp Glu Asn Ser Gln Ala Lys Pro Phe
            20   25   30

Leu Ile Asp Phe Phe Glu Val Phe Gly Ile Thr Asn Lys Arg Val Ala
        35   40   45

Thr Phe Glu His Ala Val Lys Lys Phe Ala Lys Ala His Lys Glu Gln
    50   55   60

Ser Arg Gly Phe Val Asp Leu Phe Trp Pro Gly Ile Leu Leu Ile Glu
65   70   75   80

Met Lys Ser Arg Gly Lys Asp Leu Asp Lys Ala Tyr Asp Gln Ala Leu
                85   90   95

Asp Tyr Phe Ser Gly Ile Ala Glu Arg Asp Leu Pro Arg Tyr Val Leu
            100   105   110

Val Cys Asp Phe Gln Arg Phe Arg Leu Thr Asp Leu Ile Thr Lys Glu
        115   120   125

Ser Val Glu Phe Leu Leu Lys Asp Leu Tyr Gln Asn Val Arg Ser Phe
    130   135   140

Gly Phe Ile Ala Gly Tyr Gln Thr Gln Val Ile Lys Pro Gln Asp Pro
145   150   155   160

Ile Asn Ile Lys Ala Ala Glu Arg Met Gly Lys Leu His Asp Thr Leu
                165   170   175

Lys Leu Val Gly Tyr Glu Gly His Ala Leu Glu Leu Tyr Leu Val Arg
            180   185   190

Leu Leu Phe Cys Leu Phe Ala Glu Asp Thr Thr Ile Phe Glu Lys Ser
        195   200   205

Leu Phe Gln Glu Tyr Ile Glu Thr Lys Thr Leu Glu Asp Gly Ser Asp
    210   215   220

Leu Ala His His Ile Asn Thr Leu Phe Tyr Val Leu Asn Thr Pro Glu
225   230   235   240

Gln Lys Arg Leu Lys Asn Leu Asp Glu His Leu Ala Ala Phe Pro Tyr
                245   250   255

```
Ile Asn Gly Lys Leu Phe Glu Glu Pro Leu Pro Pro Ala Gln Phe Asp
            260                 265                 270

Lys Ala Met Arg Glu Ala Leu Leu Asp Leu Cys Ser Leu Asp Trp Ser
        275                 280                 285

Arg Ile Ser Pro Ala Ile Phe Gly Ser Leu Phe Gln Ser Ile Met Asp
    290                 295                 300

Ala Lys Lys Arg Arg Asn Leu Gly Ala His Tyr Thr Ser Glu Ala Asn
305                 310                 315                 320

Ile Leu Lys Leu Ile Lys Pro Leu Phe Leu Asp Glu Leu Trp Val Glu
                325                 330                 335

Phe Glu Lys Val Lys Asn Asn Lys Asn Lys Leu Leu Ala Phe His Lys
            340                 345                 350

Lys Leu Arg Gly Leu Thr Phe Phe Asp Pro Ala Cys Gly Cys Gly Asn
        355                 360                 365

Phe Leu Val Ile Thr Tyr Arg Glu Leu Arg Leu Leu Glu Ile Glu Val
    370                 375                 380

Leu Arg Gly Leu His Arg Gly Gly Gln Gln Val Leu Asp Ile Glu His
385                 390                 395                 400

Leu Ile Gln Ile Asn Val Asp Gln Phe Phe Gly Ile Glu Ile Glu Glu
                405                 410                 415

Phe Pro Ala Gln Ile Ala Gln Val Ala Leu Trp Leu Thr Asp His Gln
            420                 425                 430

Met Asn Met Lys Ile Ser Asp Glu Phe Gly Asn Tyr Phe Ala Arg Ile
        435                 440                 445

Pro Leu Lys Ser Thr Pro His Ile Leu Asn Ala Asn Ala Leu Gln Ile
    450                 455                 460

Asp Trp Asn Asp Val Leu Glu Ala Lys Lys Cys Cys Phe Ile Leu Gly
465                 470                 475                 480

Asn Pro Pro Phe Val Gly Lys Ser Lys Gln Thr Pro Gly Gln Lys Ala
                485                 490                 495

Asp Leu Leu Ser Val Phe Gly Asn Leu Lys Ser Ala Ser Asp Leu Asp
            500                 505                 510

Leu Val Ala Ala Trp Tyr Pro Lys Ala Ala His Tyr Ile Gln Thr Asn
        515                 520                 525
```

```
Ala Asn Ile Arg Cys Ala Phe Val Ser Thr Asn Ser Ile Thr Gln Gly
    530                 535                 540

Glu Gln Val Ser Leu Leu Trp Pro Leu Leu Leu Ser Leu Gly Ile Lys
545                 550                 555                 560

Ile Asn Phe Ala His Arg Thr Phe Ser Trp Thr Asn Glu Ala Ser Gly
                565                 570                 575

Val Ala Ala Val His Cys Val Ile Ile Gly Phe Gly Leu Lys Asp Ser
            580                 585                 590

Asp Glu Lys Ile Ile Tyr Glu Tyr Glu Ser Ile Asn Gly Glu Pro Leu
        595                 600                 605

Ala Ile Lys Ala Lys Asn Ile Asn Pro Tyr Leu Arg Asp Gly Val Asp
    610                 615                 620

Val Ile Ala Cys Lys Arg Gln Gln Pro Ile Ser Lys Leu Pro Ser Met
625                 630                 635                 640

Arg Tyr Gly Asn Lys Pro Thr Asp Asp Gly Asn Phe Leu Phe Thr Asp
                645                 650                 655

Glu Glu Lys Asn Gln Phe Ile Thr Asn Glu Pro Ser Ser Glu Lys Tyr
            660                 665                 670

Phe Arg Arg Phe Val Gly Gly Asp Glu Phe Ile Asn Asn Thr Ser Arg
        675                 680                 685

Trp Cys Leu Trp Leu Asp Gly Ala Asp Ile Ser Glu Ile Arg Ala Met
    690                 695                 700

Pro Leu Val Leu Ala Arg Ile Lys Lys Val Gln Glu Phe Arg Leu Lys
705                 710                 715                 720

Ser Ser Ala Lys Pro Thr Arg Gln Ser Ala Ser Thr Pro Met Lys Phe
                725                 730                 735

Phe Tyr Ile Ser Gln Pro Asp Thr Asp Tyr Leu Leu Ile Pro Glu Thr
            740                 745                 750

Ser Ser Glu Asn Arg Gln Phe Ile Pro Ile Gly Phe Val Asp Arg Asn
        755                 760                 765

Val Ile Ser Ser Asn Ala Thr Tyr His Ile Pro Ser Ala Glu Pro Leu
    770                 775                 780

Ile Phe Gly Leu Leu Ser Ser Thr Met His Asn Cys Trp Met Arg Asn
785                 790                 795                 800
```

Val Gly Gly Arg Leu Glu Ser Arg Tyr Arg Tyr Ser Ala Ser Leu Val
805 810 815

Tyr Asn Thr Phe Pro Trp Ile Gln Pro Asn Glu Lys Gln Ser Lys Ala
820 825 830

Ile Glu Glu Ala Ala Phe Ala Ile Leu Lys Ala Arg Ser Asn Tyr Pro
835 840 845

Asn Glu Ser Leu Ala Gly Leu Tyr Asp Pro Lys Thr Met Pro Ser Glu
850 855 860

Leu Leu Lys Ala His Gln Lys Leu Asp Lys Ala Val Asp Ser Val Tyr
865 870 875 880

Gly Phe Lys Gly Pro Asn Thr Glu Ile Ala Arg Ile Ala Phe Leu Phe
885 890 895

Glu Thr Tyr Gln Lys Met Thr Ser Leu Leu Pro Pro Glu Lys Glu Ile
900 905 910

Lys Lys Ser Lys Gly Lys Asn
915

<210> 11

<211> 576
<212> PRT
<213> CEA7_ECOLX

<220>
<223> Colicin-E7 (CEA7_ECOLX) Q47112.2

<400> 11
Met Ser Gly Gly Asp Gly Arg Gly His Asn Ser Gly Ala His Asn Thr
1 5 10 15

Gly Gly Asn Ile Asn Gly Gly Pro Thr Gly Leu Gly Gly Asn Gly Gly
20 25 30

Ala Ser Asp Gly Ser Gly Trp Ser Ser Glu Asn Asn Pro Trp Gly Gly
35 40 45

Gly Ser Gly Ser Gly Val His Trp Gly Gly Gly Ser Gly His Gly Asn
50 55 60

Gly Gly Gly Asn Ser Asn Ser Gly Gly Gly Ser Asn Ser Ser Val Ala
65 70 75 80

Ala Pro Met Ala Phe Gly Phe Pro Ala Leu Ala Ala Pro Gly Ala Gly
85 90 95

Thr Leu Gly Ile Ser Val Ser Gly Glu Ala Leu Ser Ala Ala Ile Ala

100 105 110

Asp Ile Phe Ala Ala Leu Lys Gly Pro Phe Lys Phe Ser Ala Trp Gly
115 120 125

Ile Ala Leu Tyr Gly Ile Leu Pro Ser Glu Ile Ala Lys Asp Asp Pro
130 135 140

Asn Met Met Ser Lys Ile Val Thr Ser Leu Pro Ala Glu Thr Val Thr
145 150 155 160

Asn Val Gln Val Ser Thr Leu Pro Leu Asp Gln Ala Thr Val Ser Val
165 170 175

Thr Lys Arg Val Thr Asp Val Val Lys Asp Thr Arg Gln His Ile Ala
180 185 190

Val Val Ala Gly Val Pro Met Ser Val Pro Val Val Asn Ala Lys Pro
195 200 205

Thr Arg Thr Pro Gly Val Phe His Ala Ser Phe Pro Gly Val Pro Ser
210 215 220

Leu Thr Val Ser Thr Val Lys Gly Leu Pro Val Ser Thr Thr Leu Pro
225 230 235 240

Arg Gly Ile Thr Glu Asp Lys Gly Arg Thr Ala Val Pro Ala Gly Phe
245 250 255

Thr Phe Gly Gly Gly Ser His Glu Ala Val Ile Arg Phe Pro Lys Glu
260 265 270

Ser Gly Gln Lys Pro Val Tyr Val Ser Val Thr Asp Val Leu Thr Pro
275 280 285

Ala Gln Val Lys Gln Arg Gln Asp Glu Glu Lys Arg Leu Gln Gln Glu
290 295 300

Trp Asn Asp Ala His Pro Val Glu Val Ala Glu Arg Asn Tyr Glu Gln
305 310 315 320

Ala Arg Ala Glu Leu Asn Gln Ala Asn Lys Asp Val Ala Arg Asn Gln
325 330 335

Glu Arg Gln Ala Lys Ala Val Gln Val Tyr Asn Ser Arg Lys Ser Glu
340 345 350

Leu Asp Ala Ala Asn Lys Thr Leu Ala Asp Ala Lys Ala Glu Ile Lys
355 360 365

Gln Phe Glu Arg Phe Ala Arg Glu Pro Met Ala Ala Gly His Arg Met

370 375 380

Trp Gln Met Ala Gly Leu Lys Ala Gln Arg Ala Gln Thr Asp Val Asn
385 390 395 400

Asn Lys Lys Ala Ala Phe Asp Ala Ala Ala Lys Glu Lys Ser Asp Ala
405 410 415

Asp Val Ala Leu Ser Ser Ala Leu Glu Arg Arg Lys Gln Lys Glu Asn
420 425 430

Lys Glu Lys Asp Ala Lys Ala Lys Leu Asp Lys Glu Ser Lys Arg Asn
435 440 445

Lys Pro Gly Lys Ala Thr Gly Lys Gly Lys Pro Val Asn Asn Lys Trp
450 455 460

Leu Asn Asn Ala Gly Lys Asp Leu Gly Ser Pro Val Pro Asp Arg Ile
465 470 475 480

Ala Asn Lys Leu Arg Asp Lys Glu Phe Lys Ser Phe Asp Asp Phe Arg
485 490 495

Lys Lys Phe Trp Glu Glu Val Ser Lys Asp Pro Glu Leu Ser Lys Gln
500 505 510

Phe Ser Arg Asn Asn Asn Asp Arg Met Lys Val Gly Lys Ala Pro Lys
515 520 525

Thr Arg Thr Gln Asp Val Ser Gly Lys Arg Thr Ser Phe Glu Leu His
530 535 540

His Glu Lys Pro Ile Ser Gln Asn Gly Gly Val Tyr Asp Met Asp Asn
545 550 555 560

Ile Ser Val Val Thr Pro Lys Arg His Ile Asp Ile His Arg Gly Lys
565 570 575

<210> 12

<211> 274
<212> PRT
<213> Streptococcus pneumoniae

<220>
<223> End A CAA38134.1

<400> 12
Met Asn Lys Lys Thr Arg Gln Thr Leu Ile Gly Leu Leu Val Leu Leu
1 5 10 15

Leu Leu Ser Thr Gly Ser Tyr Tyr Ile Lys Gln Met Pro Ser Ala Pro
20 25 30

```
Asn Ser Pro Lys Thr Asn Leu Ser Gln Lys Lys Gln Ala Ser Glu Ala
        35                  40                  45

Pro Ser Gln Ala Leu Ala Glu Ser Val Leu Thr Asp Ala Val Lys Ser
    50                  55                  60

Gln Ile Lys Gly Ser Leu Glu Trp Asn Gly Ser Gly Ala Phe Ile Val
65                  70                  75                  80

Asn Gly Asn Lys Thr Asn Leu Asp Ala Lys Val Ser Ser Lys Pro Tyr
                85                  90                  95

Ala Asp Asn Lys Thr Lys Thr Val Gly Lys Glu Thr Val Pro Thr Val
            100                 105                 110

Ala Asn Ala Leu Leu Ser Lys Ala Thr Arg Gln Tyr Lys Asn Arg Lys
        115                 120                 125

Glu Thr Gly Asn Gly Ser Thr Ser Trp Thr Pro Pro Gly Trp His Gln
    130                 135                 140

Val Lys Asn Leu Lys Gly Ser Tyr Thr His Ala Val Asp Arg Gly His
145                 150                 155                 160

Leu Leu Gly Tyr Ala Leu Ile Gly Gly Leu Asp Gly Phe Asp Ala Ser
                165                 170                 175

Thr Ser Asn Pro Lys Asn Ile Ala Val Gln Thr Ala Trp Ala Asn Gln
            180                 185                 190

Ala Gln Ala Glu Tyr Ser Thr Gly Gln Asn Tyr Tyr Glu Ser Lys Val
        195                 200                 205

Arg Lys Ala Leu Asp Gln Asn Lys Arg Val Arg Tyr Arg Val Thr Leu
    210                 215                 220

Tyr Tyr Ala Ser Asn Glu Asp Leu Val Pro Ser Ala Ser Gln Ile Glu
225                 230                 235                 240

Ala Lys Ser Ser Asp Gly Glu Leu Glu Phe Asn Val Leu Val Pro Asn
                245                 250                 255

Val Gln Lys Gly Leu Gln Leu Asp Tyr Arg Thr Gly Glu Val Thr Val
            260                 265                 270

Thr Gln
```

<210> 13

<211> 235

<212> PRT
<213> Escherichia coli

<220>
<223> Endo I (END1_ECOLI) P25736.1

<400> 13

```
Met Tyr Arg Tyr Leu Ser Ile Ala Ala Val Val Leu Ser Ala Ala Phe
1               5                  10                  15

Ser Gly Pro Ala Leu Ala Glu Gly Ile Asn Ser Phe Ser Gln Ala Lys
            20                  25                  30

Ala Ala Ala Val Lys Val His Ala Asp Ala Pro Gly Thr Phe Tyr Cys
        35                  40                  45

Gly Cys Lys Ile Asn Trp Gln Gly Lys Lys Gly Val Val Asp Leu Gln
    50                  55                  60

Ser Cys Gly Tyr Gln Val Arg Lys Asn Glu Asn Arg Ala Ser Arg Val
65                  70                  75                  80

Glu Trp Glu His Val Val Pro Ala Trp Gln Phe Gly His Gln Arg Gln
                85                  90                  95

Cys Trp Gln Asp Gly Gly Arg Lys Asn Cys Ala Lys Asp Pro Val Tyr
            100                 105                 110

Arg Lys Met Glu Ser Asp Met His Asn Leu Gln Pro Ser Val Gly Glu
        115                 120                 125

Val Asn Gly Asp Arg Gly Asn Phe Met Tyr Ser Gln Trp Asn Gly Gly
    130                 135                 140

Glu Gly Gln Tyr Gly Gln Cys Ala Met Lys Val Asp Phe Lys Glu Lys
145                 150                 155                 160

Ala Ala Glu Pro Pro Ala Arg Ala Arg Gly Ala Ile Ala Arg Thr Tyr
                165                 170                 175

Phe Tyr Met Arg Asp Gln Tyr Asn Leu Thr Leu Ser Arg Gln Gln Thr
            180                 185                 190

Gln Leu Phe Asn Ala Trp Asn Lys Met Tyr Pro Val Thr Asp Trp Glu
        195                 200                 205

Cys Glu Arg Asp Glu Arg Ile Ala Lys Val Gln Gly Asn His Asn Pro
    210                 215                 220

Tyr Val Gln Arg Ala Cys Gln Ala Arg Lys Ser
225                 230                 235
```

<210> 14

<211> 297
<212> PRT
<213> Homo sapiens

<220>
<223> Human Endo G (NUCG_HUMAN) Q14249.4

<400> 14

```
Met Arg Ala Leu Arg Ala Gly Leu Thr Leu Ala Ser Gly Ala Gly Leu
1               5                   10                  15

Gly Ala Val Val Glu Gly Trp Arg Arg Arg Arg Glu Asp Ala Arg Ala
            20                  25                  30

Ala Pro Gly Leu Leu Gly Arg Leu Pro Val Leu Pro Val Ala Ala Ala
        35                  40                  45

Ala Glu Leu Pro Pro Val Pro Gly Gly Pro Arg Gly Pro Gly Glu Leu
    50                  55                  60

Ala Lys Tyr Gly Leu Pro Gly Leu Ala Gln Leu Lys Ser Arg Glu Ser
65                  70                  75                  80

Tyr Val Leu Cys Tyr Asp Pro Arg Thr Arg Gly Ala Leu Trp Val Val
                85                  90                  95

Glu Gln Leu Arg Pro Glu Arg Leu Arg Gly Asp Gly Asp Arg Arg Glu
            100                 105                 110

Cys Asp Phe Arg Glu Asp Asp Ser Val His Ala Tyr His Arg Ala Thr
        115                 120                 125

Asn Ala Asp Tyr Arg Gly Ser Gly Phe Asp Arg Gly His Leu Ala Ala
    130                 135                 140

Ala Ala Asn His Arg Trp Ser Gln Lys Ala Met Asp Asp Thr Phe Tyr
145                 150                 155                 160

Leu Ser Asn Val Ala Pro Gln Val Pro His Leu Asn Gln Asn Ala Trp
                165                 170                 175

Asn Asn Leu Glu Lys Tyr Ser Arg Ser Leu Thr Arg Ser Tyr Gln Asn
            180                 185                 190

Val Tyr Val Cys Thr Gly Pro Leu Phe Leu Pro Arg Thr Glu Ala Asp
        195                 200                 205

Gly Lys Ser Tyr Val Lys Tyr Gln Val Ile Gly Lys Asn His Val Ala
    210                 215                 220

Val Pro Thr His Phe Phe Lys Val Leu Ile Leu Glu Ala Ala Gly Gly
225                 230                 235                 240
```

```
Gln Ile Glu Leu Arg Thr Tyr Val Met Pro Asn Ala Pro Val Asp Glu
                245                 250                 255

Ala Ile Pro Leu Glu Arg Phe Leu Val Pro Ile Glu Ser Ile Glu Arg
            260                 265                 270

Ala Ser Gly Leu Leu Phe Val Pro Asn Ile Leu Ala Arg Ala Gly Ser
        275                 280                 285

Leu Lys Ala Ile Thr Ala Gly Ser Lys
    290                 295
```

<210> 15

<211> 299
<212> PRT
<213> Bos taurus

<220>
<223> Bovine Endo G (NUCG_BOVIN) P38447.1

<400> 15

```
Met Gln Leu Leu Arg Ala Gly Leu Thr Leu Ala Leu Gly Ala Gly Leu
1               5                   10                  15

Gly Ala Ala Ala Glu Ser Trp Trp Arg Gln Arg Ala Asp Ala Arg Ala
            20                  25                  30

Thr Pro Gly Leu Leu Ser Arg Leu Pro Val Leu Pro Val Ala Ala Ala
        35                  40                  45

Ala Gly Leu Pro Ala Val Pro Gly Ala Pro Ala Gly Gly Gly Pro Gly
    50                  55                  60

Glu Leu Ala Lys Tyr Gly Leu Pro Gly Val Ala Gln Leu Lys Ser Arg
65                  70                  75                  80

Ala Ser Tyr Val Leu Cys Tyr Asp Pro Arg Thr Arg Gly Ala Leu Trp
                85                  90                  95

Val Val Glu Gln Leu Arg Pro Glu Gly Leu Arg Gly Asp Gly Asn Arg
            100                 105                 110

Ser Ser Cys Asp Phe His Glu Asp Asp Ser Val His Ala Tyr His Arg
        115                 120                 125

Ala Thr Asn Ala Asp Tyr Arg Gly Ser Gly Phe Asp Arg Gly His Leu
    130                 135                 140

Ala Ala Ala Ala Asn His Arg Trp Ser Gln Lys Ala Met Asp Asp Thr
145                 150                 155                 160

Phe Tyr Leu Ser Asn Val Ala Pro Gln Val Pro His Leu Asn Gln Asn
```

165 170 175

Ala Trp Asn Asn Leu Glu Lys Tyr Ser Arg Ser Leu Thr Arg Thr Tyr
180 185 190

Gln Asn Val Tyr Val Cys Thr Gly Pro Leu Phe Leu Pro Arg Thr Glu
195 200 205

Ala Asp Gly Lys Ser Tyr Val Lys Tyr Gln Val Ile Gly Lys Asn His
210 215 220

Val Ala Val Pro Thr His Phe Phe Lys Val Leu Ile Leu Glu Ala Ala
225 230 235 240

Gly Gly Gln Ile Glu Leu Arg Ser Tyr Val Met Pro Asn Ala Pro Val
245 250 255

Asp Glu Ala Ile Pro Leu Glu His Phe Leu Val Pro Ile Glu Ser Ile
260 265 270

Glu Arg Ala Ser Gly Leu Leu Phe Val Pro Asn Ile Leu Ala Arg Ala
275 280 285

Gly Ser Leu Lys Ala Ile Thr Ala Gly Ser Lys
290 295

<210> 16

<211> 247
<212> PRT
<213> Haemophilus influenzae

<220>
<223> R.HinP1I AAW33811.1

<400> 16
Met Asn Leu Val Glu Leu Gly Ser Lys Thr Ala Lys Asp Gly Phe Lys
1 5 10 15

Asn Glu Lys Asp Ile Ala Asp Arg Phe Glu Asn Trp Lys Glu Asn Ser
20 25 30

Glu Ala Gln Asp Trp Leu Val Thr Met Gly His Asn Leu Asp Glu Ile
35 40 45

Lys Ser Val Lys Ala Val Val Leu Ser Gly Tyr Lys Ser Asp Ile Asn
50 55 60

Val Gln Val Leu Val Phe Tyr Lys Asp Ala Leu Asp Ile His Asn Ile
65 70 75 80

Gln Val Lys Leu Val Ser Asn Lys Arg Gly Phe Asn Gln Ile Asp Lys
85 90 95

His Trp Leu Ala His Tyr Gln Glu Met Trp Lys Phe Asp Asp Asn Leu
100 105 110

Leu Arg Ile Leu Arg His Phe Thr Gly Glu Leu Pro Pro Tyr His Ser
115 120 125

Asn Thr Lys Asp Lys Arg Arg Met Phe Met Thr Glu Phe Ser Gln Glu
130 135 140

Glu Gln Asn Ile Val Leu Asn Trp Leu Glu Lys Asn Arg Val Leu Val
145 150 155 160

Leu Thr Asp Ile Leu Arg Gly Arg Gly Asp Phe Ala Ala Glu Trp Val
165 170 175

Leu Val Ala Gln Lys Val Ser Asn Asn Ala Arg Trp Ile Leu Arg Asn
180 185 190

Ile Asn Glu Val Leu Gln His Tyr Gly Ser Gly Asp Ile Ser Leu Ser
195 200 205

Pro Arg Gly Ser Ile Asn Phe Gly Arg Val Thr Ile Gln Arg Lys Gly
210 215 220

Gly Asp Asn Gly Arg Glu Thr Ala Asn Met Leu Gln Phe Lys Ile Asp
225 230 235 240

Pro Thr Glu Leu Phe Asp Ile
245

<210> 17

<211> 188
<212> PRT
<213> Bacillus phage Bastille

<220>
<223> I-BasI AAO93095.1

<400> 17
Met Phe Gln Glu Glu Trp Lys Asp Val Thr Gly Phe Glu Asp Tyr Tyr
1 5 10 15

Glu Val Ser Asn Lys Gly Arg Val Ala Ser Lys Arg Thr Gly Val Ile
20 25 30

Met Ala Gln Tyr Lys Ile Asn Ser Gly Tyr Leu Cys Ile Lys Phe Thr
35 40 45

Val Asn Lys Lys Arg Thr Ser His Leu Val His Arg Leu Val Ala Arg
50 55 60

Glu Phe Cys Glu Gly Tyr Ser Pro Glu Leu Asp Val Asn His Lys Asp

65 70 75 80

Thr Asp Arg Met Asn Asn Asn Tyr Asp Asn Leu Glu Trp Leu Thr Arg
85 90 95

Ala Asp Asn Leu Lys Asp Val Arg Glu Arg Gly Lys Leu Asn Thr His
100 105 110

Thr Ala Arg Glu Ala Leu Ala Lys Val Ser Lys Lys Ala Val Asp Val
115 120 125

Tyr Thr Lys Asp Gly Ser Glu Tyr Ile Ala Thr Tyr Pro Ser Ala Thr
130 135 140

Glu Ala Ala Glu Ala Leu Gly Val Gln Gly Ala Lys Ile Ser Thr Val
145 150 155 160

Cys His Gly Lys Arg Gln His Thr Gly Gly Tyr His Phe Lys Phe Asn
165 170 175

Ser Ser Val Asp Pro Asn Arg Ser Val Ser Lys Lys
180 185

<210> 18

<211> 266
<212> PRT
<213> Bacillus mojavensis

<220>
<223> I-BmoI AAK09365.1

<400> 18
Met Lys Ser Gly Val Tyr Lys Ile Thr Asn Lys Asn Thr Gly Lys Phe
1 5 10 15

Tyr Ile Gly Ser Ser Glu Asp Cys Glu Ser Arg Leu Lys Val His Phe
20 25 30

Arg Asn Leu Lys Asn Asn Arg His Ile Asn Arg Tyr Leu Asn Asn Ser
35 40 45

Phe Asn Lys His Gly Glu Gln Val Phe Ile Gly Glu Val Ile His Ile
50 55 60

Leu Pro Ile Glu Glu Ala Ile Ala Lys Glu Gln Trp Tyr Ile Asp Asn
65 70 75 80

Phe Tyr Glu Glu Met Tyr Asn Ile Ser Lys Ser Ala Tyr His Gly Gly
85 90 95

Asp Leu Thr Ser Tyr His Pro Asp Lys Arg Asn Ile Ile Leu Lys Arg
100 105 110

Ala Asp Ser Leu Lys Lys Val Tyr Leu Lys Met Thr Ser Glu Glu Lys
115 120 125

Ala Lys Arg Trp Gln Cys Val Gln Gly Glu Asn Asn Pro Met Phe Gly
130 135 140

Arg Lys His Thr Glu Thr Thr Lys Leu Lys Ile Ser Asn His Asn Lys
145 150 155 160

Leu Tyr Tyr Ser Thr His Lys Asn Pro Phe Lys Gly Lys Lys His Ser
165 170 175

Glu Glu Ser Lys Thr Lys Leu Ser Glu Tyr Ala Ser Gln Arg Val Gly
180 185 190

Glu Lys Asn Pro Phe Tyr Gly Lys Thr His Ser Asp Glu Phe Lys Thr
195 200 205

Tyr Met Ser Lys Lys Phe Lys Gly Arg Lys Pro Lys Asn Ser Arg Pro
210 215 220

Val Ile Ile Asp Gly Thr Glu Tyr Glu Ser Ala Thr Glu Ala Ser Arg
225 230 235 240

Gln Leu Asn Val Val Pro Ala Thr Ile Leu His Arg Ile Lys Ser Lys
245 250 255

Asn Glu Lys Tyr Ser Gly Tyr Phe Tyr Lys
260 265

<210> 19

<211> 174
<212> PRT
<213> HMUI_BPSP1

<220>
<223> I-HmuI P34081.1

<400> 19
Met Glu Trp Lys Asp Ile Lys Gly Tyr Glu Gly His Tyr Gln Val Ser
1 5 10 15

Asn Thr Gly Glu Val Tyr Ser Ile Lys Ser Gly Lys Thr Leu Lys His
20 25 30

Gln Ile Pro Lys Asp Gly Tyr His Arg Ile Gly Leu Phe Lys Gly Gly
35 40 45

Lys Gly Lys Thr Phe Gln Val His Arg Leu Val Ala Ile His Phe Cys
50 55 60

Glu Gly Tyr Glu Glu Gly Leu Val Val Asp His Lys Asp Gly Asn Lys

65 70 75 80

Asp Asn Asn Leu Ser Thr Asn Leu Arg Trp Val Thr Gln Lys Ile Asn
85 90 95

Val Glu Asn Gln Met Ser Arg Gly Thr Leu Asn Val Ser Lys Ala Gln
100 105 110

Gln Ile Ala Lys Ile Lys Asn Gln Lys Pro Ile Ile Val Ile Ser Pro
115 120 125

Asp Gly Ile Glu Lys Glu Tyr Pro Ser Thr Lys Cys Ala Cys Glu Glu
130 135 140

Leu Gly Leu Thr Arg Gly Lys Val Thr Asp Val Leu Lys Gly His Arg
145 150 155 160

Ile His His Lys Gly Tyr Thr Phe Arg Tyr Lys Leu Asn Gly
165 170

<210> 20

<211> 245
<212> PRT
<213> TEV1_BPT4

<220>
<223> I-TevI P13299.2

<400> 20
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1 5 10 15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
20 25 30

Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
35 40 45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
50 55 60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65 70 75 80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
85 90 95

Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
100 105 110

Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly
115 120 125

Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp Asn
130 135 140

Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr Ser
145 150 155 160

Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn Ser
165 170 175

Phe Phe Asn His Lys His Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu
180 185 190

Lys Met Lys Gly Lys Lys Pro Ser Asn Ile Lys Lys Ile Ser Cys Asp
195 200 205

Gly Val Ile Phe Asp Cys Ala Ala Asp Ala Ala Arg His Phe Lys Ile
210 215 220

Ser Ser Gly Leu Val Thr Tyr Arg Val Lys Ser Asp Lys Trp Asn Trp
225 230 235 240

Phe Tyr Ile Asn Ala
245

<210> 21

<211> 258
<212> PRT
<213> TEV2_BPT4

<220>
<223> I-TevII P07072.2

<400> 21
Met Lys Trp Lys Leu Arg Lys Ser Leu Lys Ile Ala Asn Ser Val Ala
1 5 10 15

Phe Thr Tyr Met Val Arg Phe Pro Asp Lys Ser Phe Tyr Ile Gly Phe
20 25 30

Lys Lys Phe Lys Thr Ile Tyr Gly Lys Asp Thr Asn Trp Lys Glu Tyr
35 40 45

Asn Ser Ser Ser Lys Leu Val Lys Glu Lys Leu Lys Asp Tyr Lys Ala
50 55 60

Lys Trp Ile Ile Leu Gln Val Phe Asp Ser Tyr Glu Ser Ala Leu Lys
65 70 75 80

His Glu Glu Met Leu Ile Arg Lys Tyr Phe Asn Asn Glu Phe Ile Leu
85 90 95

Asn Lys Ser Ile Gly Gly Tyr Lys Phe Asn Lys Tyr Pro Asp Ser Glu

100 105 110

Glu His Lys Gln Lys Leu Ser Asn Ala His Lys Gly Lys Ile Leu Ser
115 120 125

Leu Lys His Lys Asp Lys Ile Arg Glu Lys Leu Ile Glu His Tyr Lys
130 135 140

Asn Asn Ser Arg Ser Glu Ala His Val Lys Asn Asn Ile Gly Ser Arg
145 150 155 160

Thr Ala Lys Lys Thr Val Ser Ile Ala Leu Lys Ser Gly Asn Lys Phe
165 170 175

Arg Ser Phe Lys Ser Ala Ala Lys Phe Leu Lys Cys Ser Glu Glu Gln
180 185 190

Val Ser Asn His Pro Asn Val Ile Asp Ile Lys Ile Thr Ile His Pro
195 200 205

Val Pro Glu Tyr Val Lys Ile Asn Asp Asn Ile Tyr Lys Ser Phe Val
210 215 220

Asp Ala Ala Lys Asp Leu Lys Leu His Pro Ser Arg Ile Lys Asp Leu
225 230 235 240

Cys Leu Asp Asp Asn Tyr Pro Asn Tyr Ile Val Ser Tyr Lys Arg Val
245 250 255

Glu Lys

<210> 22

<211> 269
<212> PRT
<213> TEV3_BPR03

<220>
<223> I-TevIII Q38419.1

<400> 22
Met Asn Tyr Arg Lys Ile Trp Ile Asp Ala Asn Gly Pro Ile Pro Lys
1 5 10 15

Asp Ser Asp Gly Arg Thr Asp Glu Ile His His Lys Asp Gly Asn Arg
20 25 30

Glu Asn Asn Asp Leu Asp Asn Leu Met Cys Leu Ser Ile Gln Glu His
35 40 45

Tyr Asp Ile His Leu Ala Gln Lys Asp Tyr Gln Ala Cys His Ala Ile
50 55 60

Lys Leu Arg Met Lys Tyr Ser Pro Glu Glu Ile Ser Glu Leu Ala Ser
65 70 75 80

Lys Ala Ala Lys Ser Arg Glu Ile Gln Ile Phe Asn Ile Pro Glu Val
85 90 95

Arg Ala Lys Asn Ile Ala Ser Ile Lys Ser Lys Ile Glu Asn Gly Thr
100 105 110

Phe His Leu Leu Asp Gly Glu Ile Gln Arg Lys Ser Asn Leu Asn Arg
115 120 125

Val Ala Leu Gly Ile His Asn Phe Gln Gln Ala Glu His Ile Ala Lys
130 135 140

Val Lys Glu Arg Asn Ile Ala Ala Ile Lys Glu Gly Thr His Val Phe
145 150 155 160

Cys Gly Gly Lys Met Gln Ser Glu Thr Gln Ser Lys Arg Val Asn Asp
165 170 175

Gly Ser His His Phe Leu Ser Glu Asp His Lys Lys Arg Thr Ser Ala
180 185 190

Lys Thr Leu Glu Met Val Lys Asn Gly Thr His Pro Ala Gln Lys Glu
195 200 205

Ile Thr Cys Asp Phe Cys Gly His Ile Gly Lys Gly Pro Gly Phe Tyr
210 215 220

Leu Lys His Asn Asp Arg Cys Lys Leu Asn Pro Asn Arg Ile Gln Leu
225 230 235 240

Asn Cys Pro Tyr Cys Asp Lys Lys Asp Leu Ser Pro Ser Thr Tyr Lys
245 250 255

Arg Trp His Gly Asp Asn Cys Lys Ala Arg Phe Asn Asp
260 265

<210> 23

<211> 243
<212> PRT
<213> Staphylococcus phage Twort

<220>
<223> I-TwoI AAM00817.1

<400> 23
Met Glu Glu Leu Trp Lys Glu Ile Pro Gly Phe Asn Ser Tyr Met Ile
1 5 10 15

Ser Asn Lys Gly Gln Val Tyr Ser Arg Lys Arg Asn Lys Ile Leu Ala
20 25 30

```
Leu Arg Thr Asp Lys Asn Gly Tyr Lys Arg Ile Ser Ile Phe Asn Asn
        35                  40                  45

Glu Gly Lys Arg Ile Leu Leu Gly Val His Lys Leu Val Leu Leu Gly
    50                  55                  60

Phe Lys Gly Ile Asn Thr Glu Lys Pro Ile Pro His His Lys Asn Asn
65                  70                  75                  80

Ile Lys Asp Asp Asn Arg Leu Glu Asn Leu Glu Trp Val Thr Val Ser
                85                  90                  95

Glu Asn Thr Lys His Ala Tyr Asp Ile Gly Ala Leu Lys Ser Pro Arg
            100                 105                 110

Arg Val Thr Cys Thr Leu Tyr Tyr Lys Gly Glu Pro Leu Ser Cys Tyr
        115                 120                 125

Asp Ser Leu Phe Asp Leu Ala Lys Ala Leu Lys Val Ser Arg Ser Val
    130                 135                 140

Ile Glu Ser Pro Arg Asn Gly Leu Val Leu Ser Thr Phe Glu Val Lys
145                 150                 155                 160

Arg Glu Pro Thr Ile Gln Gly Leu Pro Leu Asn Lys Glu Ile Phe Glu
                165                 170                 175

His Ser Leu Ile Lys Gly Leu Gly Asn Pro Pro Leu Lys Val Tyr Asn
            180                 185                 190

Glu Asp Glu Thr Tyr Tyr Phe Leu Thr Leu Met Asp Ile Ser Lys Tyr
        195                 200                 205

Phe Asn Glu Ser Tyr Ser Lys Val Gln Arg Gly Tyr Tyr Lys Gly Lys
    210                 215                 220

Trp Lys Ser Tyr Ile Ile Glu His Ile Asp Phe Tyr Glu Tyr Tyr Lys
225                 230                 235                 240

Gln Thr His
```

<210> 24

<211> 262
<212> PRT
<213> T2M1_MORSP

<220>
<223> R.MspI P11405.1

<400> 24

```
Met Arg Thr Glu Leu Leu Ser Lys Leu Tyr Asp Asp Phe Gly Ile Asp
```

1 5 10 15

Gln Leu Pro His Thr Gln His Gly Val Thr Ser Asp Arg Leu Gly Lys
20 25 30

Leu Tyr Glu Lys Tyr Ile Leu Asp Ile Phe Lys Asp Ile Glu Ser Leu
35 40 45

Lys Lys Tyr Asn Thr Asn Ala Phe Pro Gln Glu Lys Asp Ile Ser Ser
50 55 60

Lys Leu Leu Lys Ala Leu Asn Leu Asp Leu Asp Asn Ile Ile Asp Val
65 70 75 80

Ser Ser Ser Asp Thr Asp Leu Gly Arg Thr Ile Ala Gly Gly Ser Pro
85 90 95

Lys Thr Asp Ala Thr Ile Arg Phe Thr Phe His Asn Gln Ser Ser Arg
100 105 110

Leu Val Pro Leu Asn Ile Lys His Ser Ser Lys Lys Lys Val Ser Ile
115 120 125

Ala Glu Tyr Asp Val Glu Thr Ile Cys Thr Gly Val Gly Ile Ser Asp
130 135 140

Gly Glu Leu Lys Glu Leu Ile Arg Lys His Gln Asn Asp Gln Ser Ala
145 150 155 160

Lys Leu Phe Thr Pro Val Gln Lys Gln Arg Leu Thr Glu Leu Leu Glu
165 170 175

Pro Tyr Arg Glu Arg Phe Ile Arg Trp Cys Val Thr Leu Arg Ala Glu
180 185 190

Lys Ser Glu Gly Asn Ile Leu His Pro Asp Leu Leu Ile Arg Phe Gln
195 200 205

Val Ile Asp Arg Glu Tyr Val Asp Val Thr Ile Lys Asn Ile Asp Asp
210 215 220

Tyr Val Ser Asp Arg Ile Ala Glu Gly Ser Lys Ala Arg Lys Pro Gly
225 230 235 240

Phe Gly Thr Gly Leu Asn Trp Thr Tyr Ala Ser Gly Ser Lys Ala Lys
245 250 255

Lys Met Gln Phe Lys Gly
260

<210> 25

<211> 246
<212> PRT
<213> Kocuria varians

<220>
<223> R.MvaI

<400> 25
Met Ser Glu Tyr Leu Asn Leu Leu Lys Glu Ala Ile Gln Asn Val Val
1 5 10 15

Asp Gly Gly Trp His Glu Thr Lys Arg Lys Gly Asn Thr Gly Ile Gly
20 25 30

Lys Thr Phe Glu Asp Leu Leu Glu Lys Glu Glu Asp Asn Leu Asp Ala
35 40 45

Pro Asp Phe His Asp Ile Glu Ile Lys Thr His Glu Thr Ala Ala Lys
50 55 60

Ser Leu Leu Thr Leu Phe Thr Lys Ser Pro Thr Asn Pro Arg Gly Ala
65 70 75 80

Asn Thr Met Leu Arg Asn Arg Tyr Gly Lys Lys Asp Glu Tyr Gly Asn
85 90 95

Asn Ile Leu His Gln Thr Val Ser Gly Asn Arg Lys Thr Asn Ser Asn
100 105 110

Ser Tyr Asn Tyr Asp Phe Lys Ile Asp Ile Asp Trp Glu Ser Gln Val
115 120 125

Val Arg Leu Glu Val Phe Asp Lys Gln Asp Ile Met Ile Asp Asn Ser
130 135 140

Val Tyr Trp Ser Phe Asp Ser Leu Gln Asn Gln Leu Asp Lys Lys Leu
145 150 155 160

Lys Tyr Ile Ala Val Ile Ser Ala Glu Ser Lys Ile Glu Asn Glu Lys
165 170 175

Lys Tyr Tyr Lys Tyr Asn Ser Ala Asn Leu Phe Thr Asp Leu Thr Val
180 185 190

Gln Ser Leu Cys Arg Gly Ile Glu Asn Gly Asp Ile Lys Val Asp Ile
195 200 205

Arg Ile Gly Ala Tyr His Ser Gly Lys Lys Lys Gly Lys Thr His Asp
210 215 220

His Gly Thr Ala Phe Arg Ile Asn Met Glu Lys Leu Leu Glu Tyr Gly
225 230 235 240

Glu Val Lys Val Ile Val
245

<210> 26

<211> 274
<212> PRT
<213> Nostoc sp. PCC 7120

<220>
<223> NucA CAA45962.1

<400> 26
Met Gly Ile Cys Gly Lys Leu Gly Val Ala Ala Leu Val Ala Leu Ile
1 5 10 15

Val Gly Cys Ser Pro Val Gln Ser Gln Val Pro Pro Leu Thr Glu Leu
20 25 30

Ser Pro Ser Ile Ser Val His Leu Leu Leu Gly Asn Pro Ser Gly Ala
35 40 45

Thr Pro Thr Lys Leu Thr Pro Asp Asn Tyr Leu Met Val Lys Asn Gln
50 55 60

Tyr Ala Leu Ser Tyr Asn Asn Ser Lys Gly Thr Ala Asn Trp Val Ala
65 70 75 80

Trp Gln Leu Asn Ser Ser Trp Leu Gly Asn Ala Glu Arg Gln Asp Asn
85 90 95

Phe Arg Pro Asp Lys Thr Leu Pro Ala Gly Trp Val Arg Val Thr Pro
100 105 110

Ser Met Tyr Ser Gly Ser Gly Tyr Asp Arg Gly His Ile Ala Pro Ser
115 120 125

Ala Asp Arg Thr Lys Thr Thr Glu Asp Asn Ala Ala Thr Phe Leu Met
130 135 140

Thr Asn Met Met Pro Gln Thr Pro Asp Asn Asn Arg Asn Thr Trp Gly
145 150 155 160

Asn Leu Glu Asp Tyr Cys Arg Glu Leu Val Ser Gln Gly Lys Glu Leu
165 170 175

Tyr Ile Val Ala Gly Pro Asn Gly Ser Leu Gly Lys Pro Leu Lys Gly
180 185 190

Lys Val Thr Val Pro Lys Ser Thr Trp Lys Ile Val Val Val Leu Asp
195 200 205

Ser Pro Gly Ser Gly Leu Glu Gly Ile Thr Ala Asn Thr Arg Val Ile

210 215 220

Ala Val Asn Ile Pro Asn Asp Pro Glu Leu Asn Asn Asp Trp Arg Ala
225 230 235 240

Tyr Lys Val Ser Val Asp Glu Leu Glu Ser Leu Thr Gly Tyr Asp Phe
245 250 255

Leu Ser Asn Val Ser Pro Asn Ile Gln Thr Ser Ile Glu Ser Lys Val
260 265 270

Asp Asn

<210> 27

<211> 232
<212> PRT
<213> NUCM_DICD3

<220>
<223> NucM P37994.2

<400> 27
Met Leu Arg Asn Leu Val Ile Phe Ala Val Leu Gly Ala Gly Leu Thr
1 5 10 15

Thr Leu Ala Ala Ala Gly Gln Asp Ile Asn Asn Phe Thr Gln Ala Lys
20 25 30

Ala Ala Ala Ala Lys Ile His Gln Asp Ala Pro Gly Thr Phe Tyr Cys
35 40 45

Gly Cys Lys Ile Asn Trp Gln Gly Lys Lys Gly Thr Pro Asp Leu Ala
50 55 60

Ser Cys Gly Tyr Gln Val Arg Lys Asp Ala Asn Arg Ala Ser Arg Ile
65 70 75 80

Glu Trp Glu His Val Val Pro Ala Trp Gln Phe Gly His Gln Arg Gln
85 90 95

Cys Trp Gln Asp Gly Gly Arg Lys Asn Cys Thr Lys Asp Asp Val Tyr
100 105 110

Arg Gln Ile Glu Thr Asp Leu His Asn Leu Gln Pro Ala Ile Gly Glu
115 120 125

Val Asn Gly Asp Arg Gly Asn Phe Met Tyr Ser Gln Trp Asn Gly Gly
130 135 140

Glu Arg Gln Tyr Gly Gln Cys Glu Met Lys Ile Asp Phe Lys Ser Gln
145 150 155 160

Leu Ala Glu Pro Pro Glu Arg Ala Arg Gly Ala Ile Ala Arg Thr Tyr
165 170 175

Phe Tyr Met Arg Asp Arg Tyr Asn Leu Asn Leu Ser Arg Gln Gln Thr
180 185 190

Gln Leu Phe Asp Ala Trp Asn Lys Gln Tyr Pro Ala Thr Thr Trp Glu
195 200 205

Cys Thr Arg Glu Lys Arg Ile Ala Ala Val Gln Gly Asn His Asn Pro
210 215 220

Tyr Val Gln Gln Ala Cys Gln Pro
225 230

<210> 28

<211> 231
<212> PRT
<213> Vibrio vulnificus

<220>
<223> Vvn AAF19759.1

<400> 28
Met Lys Arg Leu Phe Ile Phe Ile Ala Ser Phe Thr Ala Phe Ala Ile
1 5 10 15

Gln Ala Ala Pro Pro Ser Ser Phe Ser Ala Ala Lys Gln Gln Ala Val
20 25 30

Lys Ile Tyr Gln Asp His Pro Ile Ser Phe Tyr Cys Gly Cys Asp Ile
35 40 45

Glu Trp Gln Gly Lys Lys Gly Ile Pro Asn Leu Glu Thr Cys Gly Tyr
50 55 60

Gln Val Arg Lys Gln Gln Thr Arg Ala Ser Arg Ile Glu Trp Glu His
65 70 75 80

Val Val Pro Ala Trp Gln Phe Gly His His Arg Gln Cys Trp Gln Lys
85 90 95

Gly Gly Arg Lys Asn Cys Ser Lys Asn Asp Gln Gln Phe Arg Leu Met
100 105 110

Glu Ala Asp Leu His Asn Leu Thr Pro Ala Ile Gly Glu Val Asn Gly
115 120 125

Asp Arg Ser Asn Phe Asn Phe Ser Gln Trp Asn Gly Val Asp Gly Val
130 135 140

Ser Tyr Gly Arg Cys Glu Met Gln Val Asn Phe Lys Gln Arg Lys Val
145 150 155 160

Met Pro Gln Thr Glu Leu Arg Gly Ser Ile Ala Arg Thr Tyr Leu Tyr
165 170 175

Met Ser Gln Glu Tyr Gly Phe Gln Leu Ser Lys Gln Gln Gln Gln Leu
180 185 190

Met Gln Ala Trp Asn Lys Ser Tyr Pro Val Asp Glu Trp Glu Cys Thr
195 200 205

Arg Asp Asp Arg Ile Ala Lys Ile Gln Gly Asn His Asn Pro Phe Val
210 215 220

Gln Gln Ser Cys Gln Thr Gln
225 230

<210> 29

<211> 222
<212> PRT
<213> artificial sequence

<220>
<223> Vvn_CLS

<400> 29
Met Ala Ser Gly Ala Pro Pro Ser Ser Phe Ser Ala Ala Lys Gln Gln
1 5 10 15

Ala Val Lys Ile Tyr Gln Asp His Pro Ile Ser Phe Tyr Cys Gly Cys
20 25 30

Asp Ile Glu Trp Gln Gly Lys Lys Gly Ile Pro Asn Leu Glu Thr Cys
35 40 45

Gly Tyr Gln Val Arg Lys Gln Gln Thr Arg Ala Ser Arg Ile Glu Trp
50 55 60

Glu His Val Val Pro Ala Trp Gln Phe Gly His His Arg Gln Cys Trp
65 70 75 80

Gln Lys Gly Gly Arg Lys Asn Cys Ser Lys Asn Asp Gln Gln Phe Arg
85 90 95

Leu Met Glu Ala Asp Leu His Asn Leu Thr Pro Ala Ile Gly Glu Val
100 105 110

Asn Gly Asp Arg Ser Asn Phe Asn Phe Ser Gln Trp Asn Gly Val Asp
115 120 125

Gly Val Ser Tyr Gly Arg Cys Glu Met Gln Val Asn Phe Lys Gln Arg
130 135 140

```
Lys Val Met Pro Pro Asp Arg Ala Arg Gly Ser Ile Ala Arg Thr Tyr
145                 150                 155                 160


Leu Tyr Met Ser Gln Glu Tyr Gly Phe Gln Leu Ser Lys Gln Gln Gln
                165                 170                 175


Gln Leu Met Gln Ala Trp Asn Lys Ser Tyr Pro Val Asp Glu Trp Glu
            180                 185                 190


Cys Thr Arg Asp Asp Arg Ile Ala Lys Ile Gln Gly Asn His Asn Pro
        195                 200                 205


Phe Val Gln Gln Ser Cys Gln Thr Gln Gly Ser Ser Ala Asp
    210                 215                 220


<210>  30

<211>  231
<212>  PRT
<213>  NUC_STAAU

<220>
<223>  Staphylococcal nuclease (NUC_STAAU) P00644.1

<400>  30
Met Leu Val Met Thr Glu Tyr Leu Leu Ser Ala Gly Ile Cys Met Ala
1               5                   10                  15


Ile Val Ser Ile Leu Leu Ile Gly Met Ala Ile Ser Asn Val Ser Lys
            20                  25                  30


Gly Gln Tyr Ala Lys Arg Phe Phe Phe Phe Ala Thr Ser Cys Leu Val
        35                  40                  45


Leu Thr Leu Val Val Val Ser Ser Leu Ser Ser Ser Ala Asn Ala Ser
    50                  55                  60


Gln Thr Asp Asn Gly Val Asn Arg Ser Gly Ser Glu Asp Pro Thr Val
65                  70                  75                  80


Tyr Ser Ala Thr Ser Thr Lys Lys Leu His Lys Glu Pro Ala Thr Leu
                85                  90                  95


Ile Lys Ala Ile Asp Gly Asp Thr Val Lys Leu Met Tyr Lys Gly Gln
            100                 105                 110


Pro Met Thr Phe Arg Leu Leu Leu Val Asp Thr Pro Glu Thr Lys His
        115                 120                 125


Pro Lys Lys Gly Val Glu Lys Tyr Gly Pro Glu Ala Ser Ala Phe Thr
    130                 135                 140


Lys Lys Met Val Glu Asn Ala Lys Lys Ile Glu Val Glu Phe Asp Lys
145                 150                 155                 160
```

```
Gly Gln Arg Thr Asp Lys Tyr Gly Arg Gly Leu Ala Tyr Ile Tyr Ala
                165                 170                 175

Asp Gly Lys Met Val Asn Glu Ala Leu Val Arg Gln Gly Leu Ala Lys
            180                 185                 190

Val Ala Tyr Val Tyr Lys Pro Asn Asn Thr His Glu Gln His Leu Arg
        195                 200                 205

Lys Ser Glu Ala Gln Ala Lys Lys Glu Lys Leu Asn Ile Trp Ser Glu
    210                 215                 220

Asp Asn Ala Asp Ser Gly Gln
225                 230
```

<210> 31

<211> 169
<212> PRT
<213> NUC_STAHY

<220>
<223> Staphylococcal nuclease (NUC_STAHY) P43270.1

<400> 31

```
Met Lys Lys Ile Thr Thr Gly Leu Ile Ile Val Val Ala Ala Ile Ile
1               5                   10                  15

Val Leu Ser Ile Gln Phe Met Thr Glu Ser Gly Pro Phe Lys Ser Ala
            20                  25                  30

Gly Leu Ser Asn Ala Asn Glu Gln Thr Tyr Lys Val Ile Arg Val Ile
        35                  40                  45

Asp Gly Asp Thr Ile Ile Val Asp Lys Asp Gly Lys Gln Gln Asn Leu
    50                  55                  60

Arg Met Ile Gly Val Asp Thr Pro Glu Thr Val Lys Pro Asn Thr Pro
65                  70                  75                  80

Val Gln Pro Tyr Gly Lys Glu Ala Ser Asp Phe Thr Lys Arg His Leu
                85                  90                  95

Thr Asn Gln Lys Val Arg Leu Glu Tyr Asp Lys Gln Glu Lys Asp Arg
            100                 105                 110

Tyr Gly Arg Thr Leu Ala Tyr Val Trp Leu Gly Lys Glu Met Phe Asn
        115                 120                 125

Glu Lys Leu Ala Lys Glu Gly Leu Ala Arg Ala Lys Phe Tyr Arg Pro
    130                 135                 140
```

Asn Tyr Lys Tyr Gln Glu Arg Ile Glu Gln Ala Gln Lys Gln Ala Gln
145 150 155 160

Lys Leu Lys Lys Asn Ile Trp Ser Asn
165

<210> 32

<211> 174
<212> PRT
<213> NUC_SHIFL

<220>
<223> Micrococcal nuclease (NUC_SHIFL)P29769.1

<400> 32
Met Lys Ser Ala Leu Ala Ala Leu Arg Ala Val Ala Ala Ala Val Val
1 5 10 15

Leu Ile Val Ser Val Pro Ala Trp Ala Asp Phe Arg Gly Glu Val Val
20 25 30

Arg Ile Leu Asp Gly Asp Thr Ile Asp Val Leu Val Asn Arg Gln Thr
35 40 45

Ile Arg Val Arg Leu Ala Asp Ile Asp Ala Pro Glu Ser Gly Gln Ala
50 55 60

Phe Gly Ser Arg Ala Arg Gln Arg Leu Ala Asp Leu Thr Phe Arg Gln
65 70 75 80

Glu Val Gln Val Thr Glu Lys Glu Val Asp Arg Tyr Gly Arg Thr Leu
85 90 95

Gly Val Val Tyr Ala Pro Leu Gln Tyr Pro Gly Gly Gln Thr Gln Leu
100 105 110

Thr Asn Ile Asn Ala Ile Met Val Gln Glu Gly Met Ala Trp Ala Tyr
115 120 125

Arg Tyr Tyr Gly Lys Pro Thr Asp Ala Gln Met Tyr Glu Tyr Glu Lys
130 135 140

Glu Ala Arg Arg Gln Arg Leu Gly Leu Trp Ser Asp Pro Asn Ala Gln
145 150 155 160

Glu Pro Trp Lys Trp Arg Arg Ala Ser Lys Asn Ala Thr Asn
165 170

<210> 33

<211> 211
<212> PRT
<213> YNCB_BACSU

<220>
<223> Endonuclease yncB P94492.1

<400> 33
Met Lys Lys Ile Leu Ile Ser Met Ile Ala Ile Val Leu Ser Ile Thr
1 5 10 15

Leu Ala Ala Cys Gly Ser Asn His Ala Ala Lys Asn His Ser Asp Ser
20 25 30

Asn Gly Thr Glu Gln Val Ser Gln Asp Thr His Ser Asn Glu Tyr Asn
35 40 45

Gln Thr Glu Gln Lys Ala Gly Thr Pro His Ser Lys Asn Gln Lys Lys
50 55 60

Leu Val Asn Val Thr Leu Asp Arg Ala Ile Asp Gly Asp Thr Ile Lys
65 70 75 80

Val Ile Tyr Asn Gly Lys Lys Asp Thr Val Arg Tyr Leu Leu Val Asp
85 90 95

Thr Pro Glu Thr Lys Lys Pro Asn Ser Cys Val Gln Pro Tyr Gly Glu
100 105 110

Asp Ala Ser Lys Arg Asn Lys Glu Leu Val Asn Ser Gly Lys Leu Gln
115 120 125

Leu Glu Phe Asp Lys Gly Asp Arg Arg Asp Lys Tyr Gly Arg Leu Leu
130 135 140

Ala Tyr Val Tyr Val Asp Gly Lys Ser Val Gln Glu Thr Leu Leu Lys
145 150 155 160

Glu Gly Leu Ala Arg Val Ala Tyr Val Tyr Glu Pro Asn Thr Lys Tyr
165 170 175

Ile Asp Gln Phe Arg Leu Asp Glu Gln Glu Ala Lys Ser Asp Lys Leu
180 185 190

Ser Ile Trp Ser Lys Ser Gly Tyr Val Thr Asn Arg Gly Phe Asn Gly
195 200 205

Cys Val Lys
210

<210> 34

<211> 149
<212> PRT
<213> ENRN_BPT7

<220>
<223> Endodeoxyribonuclease I (ENRN_BPT7)P00641.1

<400> 34
Met Ala Gly Tyr Gly Ala Lys Gly Ile Arg Lys Val Gly Ala Phe Arg
1 5 10 15

Ser Gly Leu Glu Asp Lys Val Ser Lys Gln Leu Glu Ser Lys Gly Ile
20 25 30

Lys Phe Glu Tyr Glu Glu Trp Lys Val Pro Tyr Val Ile Pro Ala Ser
35 40 45

Asn His Thr Tyr Thr Pro Asp Phe Leu Leu Pro Asn Gly Ile Phe Val
50 55 60

Glu Thr Lys Gly Leu Trp Glu Ser Asp Asp Arg Lys Lys His Leu Leu
65 70 75 80

Ile Arg Glu Gln His Pro Glu Leu Asp Ile Arg Ile Val Phe Ser Ser
85 90 95

Ser Arg Thr Lys Leu Tyr Lys Gly Ser Pro Thr Ser Tyr Gly Glu Phe
100 105 110

Cys Glu Lys His Gly Ile Lys Phe Ala Asp Lys Leu Ile Pro Ala Glu
115 120 125

Trp Ile Lys Glu Pro Lys Lys Glu Val Pro Phe Asp Arg Leu Lys Arg
130 135 140

Lys Gly Gly Lys Lys
145

<210> 35

<211> 671
<212> PRT
<213> Homo sapiens

<220>
<223> Metnase Q53H47.1

<400> 35
Met Ala Glu Phe Lys Glu Lys Pro Glu Ala Pro Thr Glu Gln Leu Asp
1 5 10 15

Val Ala Cys Gly Gln Glu Asn Leu Pro Val Gly Ala Trp Pro Pro Gly
20 25 30

Ala Ala Pro Ala Pro Phe Gln Tyr Thr Pro Asp His Val Val Gly Pro
35 40 45

Gly Ala Asp Ile Asp Pro Thr Gln Ile Thr Phe Pro Gly Cys Ile Cys
50 55 60

```
Val Lys Thr Pro Cys Leu Pro Gly Thr Cys Ser Cys Leu Arg His Gly
65                  70                  75                  80

Glu Asn Tyr Asp Asp Asn Ser Cys Leu Arg Asp Ile Gly Ser Gly Gly
                85                  90                  95

Lys Tyr Ala Glu Pro Val Phe Glu Cys Asn Val Leu Cys Arg Cys Ser
            100                 105                 110

Asp His Cys Arg Asn Arg Val Val Gln Lys Gly Leu Gln Phe His Phe
        115                 120                 125

Gln Val Phe Lys Thr His Lys Lys Gly Trp Gly Leu Arg Thr Leu Glu
    130                 135                 140

Phe Ile Pro Lys Gly Arg Phe Val Cys Glu Tyr Ala Gly Glu Val Leu
145                 150                 155                 160

Gly Phe Ser Glu Val Gln Arg Arg Ile His Leu Gln Thr Lys Ser Asp
                165                 170                 175

Ser Asn Tyr Ile Ile Ala Ile Arg Glu His Val Tyr Asn Gly Gln Val
            180                 185                 190

Met Glu Thr Phe Val Asp Pro Thr Tyr Ile Gly Asn Ile Gly Arg Phe
        195                 200                 205

Leu Asn His Ser Cys Glu Pro Asn Leu Leu Met Ile Pro Val Arg Ile
    210                 215                 220

Asp Ser Met Val Pro Lys Leu Ala Leu Phe Ala Ala Lys Asp Ile Val
225                 230                 235                 240

Pro Glu Glu Glu Leu Ser Tyr Asp Tyr Ser Gly Arg Tyr Leu Asn Leu
                245                 250                 255

Thr Val Ser Glu Asp Lys Glu Arg Leu Asp His Gly Lys Leu Arg Lys
            260                 265                 270

Pro Cys Tyr Cys Gly Ala Lys Ser Cys Thr Ala Phe Leu Pro Phe Asp
        275                 280                 285

Ser Ser Leu Tyr Cys Pro Val Glu Lys Ser Asn Ile Ser Cys Gly Asn
    290                 295                 300

Glu Lys Glu Pro Ser Met Cys Gly Ser Ala Pro Ser Val Phe Pro Ser
305                 310                 315                 320

Cys Lys Arg Leu Thr Leu Glu Thr Met Lys Met Met Leu Asp Lys Lys
                325                 330                 335
```

```
Gln Ile Arg Ala Ile Phe Leu Phe Glu Phe Lys Met Gly Arg Lys Ala
            340                 345                 350

Ala Glu Thr Thr Arg Asn Ile Asn Asn Ala Phe Gly Pro Gly Thr Ala
        355                 360                 365

Asn Glu Arg Thr Val Gln Trp Trp Phe Lys Lys Phe Cys Lys Gly Asp
    370                 375                 380

Glu Ser Leu Glu Asp Glu Glu Arg Ser Gly Arg Pro Ser Glu Val Asp
385                 390                 395                 400

Asn Asp Gln Leu Arg Ala Ile Ile Glu Ala Asp Pro Leu Thr Thr Thr
                405                 410                 415

Arg Glu Val Ala Glu Glu Leu Asn Val Asn His Ser Thr Val Val Arg
            420                 425                 430

His Leu Lys Gln Ile Gly Lys Val Lys Lys Leu Asp Lys Trp Val Pro
        435                 440                 445

His Glu Leu Thr Glu Asn Gln Lys Asn Arg Arg Phe Glu Val Ser Ser
    450                 455                 460

Ser Leu Ile Leu Arg Asn His Asn Glu Pro Phe Leu Asp Arg Ile Val
465                 470                 475                 480

Thr Cys Asp Glu Lys Trp Ile Leu Tyr Asp Asn Arg Arg Arg Ser Ala
                485                 490                 495

Gln Trp Leu Asp Gln Glu Glu Ala Pro Lys His Phe Pro Lys Pro Ile
            500                 505                 510

Leu His Pro Lys Lys Val Met Val Thr Ile Trp Trp Ser Ala Ala Gly
        515                 520                 525

Leu Ile His Tyr Ser Phe Leu Asn Pro Gly Glu Thr Ile Thr Ser Glu
    530                 535                 540

Lys Tyr Ala Gln Glu Ile Asp Glu Met Asn Gln Lys Leu Gln Arg Leu
545                 550                 555                 560

Gln Leu Ala Leu Val Asn Arg Lys Gly Pro Ile Leu Leu His Asp Asn
                565                 570                 575

Ala Arg Pro His Val Ala Gln Pro Thr Leu Gln Lys Leu Asn Glu Leu
            580                 585                 590

Gly Tyr Glu Val Leu Pro His Pro Pro Tyr Ser Pro Asp Leu Leu Pro
        595                 600                 605
```

Thr Asn Tyr His Val Phe Lys His Leu Asn Asn Phe Leu Gln Gly Lys
610 615 620

Arg Phe His Asn Gln Gln Asp Ala Glu Asn Ala Phe Gln Glu Phe Val
625 630 635 640

Glu Ser Gln Ser Thr Asp Phe Tyr Ala Thr Gly Ile Asn Gln Leu Ile
645 650 655

Ser Arg Trp Gln Lys Cys Val Asp Cys Asn Gly Ser Tyr Phe Asp
660 665 670

<210> 36

<211> 488
<212> PRT
<213> Geobacillus stearothermophilus

<220>
<223> Nb.BsrDI ABD15132.1

<400> 36
Met Thr Glu Tyr Asp Leu His Leu Tyr Ala Asp Ser Phe His Glu Gly
1 5 10 15

His Trp Cys Cys Glu Asn Leu Ala Lys Ile Ala Gln Ser Asp Gly Gly
20 25 30

Lys His Gln Ile Asp Tyr Leu Gln Gly Phe Ile Pro Arg His Ser Leu
35 40 45

Ile Phe Ser Asp Leu Ile Ile Asn Ile Thr Val Phe Gly Ser Tyr Lys
50 55 60

Ser Trp Lys His Leu Pro Lys Gln Ile Lys Asp Leu Leu Phe Trp Gly
65 70 75 80

Lys Pro Asp Phe Ile Ala Tyr Asp Pro Lys Asn Asp Lys Ile Leu Phe
85 90 95

Ala Val Glu Glu Thr Gly Ala Val Pro Thr Gly Asn Gln Ala Leu Gln
100 105 110

Arg Cys Glu Arg Ile Tyr Gly Ser Ala Arg Lys Gln Ile Pro Phe Trp
115 120 125

Tyr Leu Leu Ser Glu Phe Gly Gln His Lys Asp Gly Gly Thr Arg Arg
130 135 140

Asp Ser Ile Trp Pro Thr Ile Met Gly Leu Lys Leu Thr Gln Leu Val
145 150 155 160

Lys Thr Pro Ser Ile Ile Leu His Tyr Ser Asp Ile Asn Asn Pro Glu
165 170 175

```
Asp Tyr Asn Ser Gly Asn Gly Leu Lys Phe Leu Phe Lys Ser Leu Leu
            180                 185                 190

Gln Ile Ile Ile Asn Tyr Cys Thr Leu Lys Asn Pro Leu Lys Gly Met
        195                 200                 205

Leu Glu Leu Leu Ser Ile Gln Tyr Glu Asn Met Leu Glu Phe Ile Lys
    210                 215                 220

Ser Gln Trp Lys Glu Gln Ile Asp Phe Leu Pro Gly Glu Glu Ile Leu
225                 230                 235                 240

Asn Thr Lys Thr Lys Glu Leu Ala Arg Met Tyr Ala Ser Leu Ala Ile
                245                 250                 255

Gly Gln Thr Val Lys Ile Pro Glu Glu Leu Phe Asn Trp Pro Arg Thr
            260                 265                 270

Asp Lys Val Asn Phe Lys Ser Pro Gln Gly Leu Ile Lys Tyr Asp Glu
        275                 280                 285

Leu Cys Tyr Gln Leu Glu Lys Ala Val Gly Ser Lys Lys Ala Tyr Cys
    290                 295                 300

Leu Ser Asn Asn Ala Gly Ala Lys Pro Gln Lys Leu Glu Ser Leu Lys
305                 310                 315                 320

Glu Trp Ile Asn Ser Gln Lys Lys Leu Phe Asp Lys Ala Pro Lys Leu
                325                 330                 335

Thr Pro Pro Ala Glu Phe Asn Met Lys Leu Asp Ala Phe Pro Val Thr
            340                 345                 350

Ser Asn Asn Asn Tyr Tyr Val Thr Thr Ser Lys Asn Ile Leu Tyr Leu
        355                 360                 365

Phe Asp Tyr Trp Lys Asp Leu Arg Ile Ala Ile Glu Thr Ala Phe Pro
    370                 375                 380

Arg Leu Lys Gly Lys Leu Pro Thr Asp Ile Asp Glu Lys Pro Ala Leu
385                 390                 395                 400

Ile Tyr Ile Cys Asn Ser Val Lys Pro Gly Arg Leu Phe Gly Asp Pro
                405                 410                 415

Phe Thr Gly Gln Leu Ser Ala Phe Ser Thr Ile Phe Gly Lys Lys Asn
            420                 425                 430

Ile Asp Met Pro Arg Ile Val Val Ala Tyr Tyr Pro His Gln Ile Tyr
        435                 440                 445
```

Ser Gln Ala Leu Pro Lys Asn Asn Lys Ser Asn Lys Gly Ile Thr Leu
450 455 460

Lys Lys Glu Leu Thr Asp Phe Leu Ile Phe His Gly Gly Val Val Val
465 470 475 480

Lys Leu Asn Glu Gly Lys Ala Tyr
485

<210> 37

<211> 217
<212> PRT
<213> Geobacillus stearothermophilus

<220>
<223> BsrDI A ABD15133.1

<400> 37
Met Thr Asp Tyr Arg Tyr Ser Phe Glu Leu Ser Glu Glu Ile Ala Arg
1 5 10 15

Trp Ala Phe Glu Ile Lys Thr Lys Asn Thr Asp Trp Phe Val Ala Phe
20 25 30

Ser Asn Pro Thr Ala Gly Pro Trp Lys Arg Val Met Ala Ile Asp Lys
35 40 45

Ala Ser Asn Arg Glu Gly Glu Val His Arg Phe Gly Arg Glu Asp Glu
50 55 60

Arg Pro Asp Ile Ile Leu Val Asn Asp Asn Ile Ser Leu Ile Leu Ile
65 70 75 80

Leu Glu Ala Lys Glu Lys Leu Asn Gln Leu Ile Ser Lys Ser Gln Val
85 90 95

Asp Lys Ser Val Asp Val Phe Leu Thr Leu Ser Ser Ile Leu Lys Glu
100 105 110

Lys Ser Asp Asn Asn Tyr Trp Gly Asp Arg Thr Lys Tyr Ile Asn Val
115 120 125

Leu Gly Ile Leu Trp Gly Ser Glu Gln Glu Thr Ser Gln Lys Asp Ile
130 135 140

Asp Asn Ala Phe Arg Val Tyr Arg Asp Ser Leu Val Lys Asn Leu Lys
145 150 155 160

Glu Ile Asn Pro Thr Pro Thr Asn Ile Cys Thr Asp Ile Leu Val Gly
165 170 175

Val Glu Ser Ile Lys Asn Lys Lys Glu Glu Ile Ser Ile Lys Ile His
180 185 190

Val Ser Asn Ile Tyr Ala Glu Ile Tyr Pro Lys Phe Thr Gly Lys His
195 200 205

Leu Leu Glu Lys Leu Ala Val Leu Asn
210 215

<210> 38

<211> 604
<212> PRT
<213> Bacillus sp. D6

<220>
<223> Nt.BspD6I ABN42182.1 (R.BspD6I large subunit)

<400> 38
Met Ala Lys Lys Val Asn Trp Tyr Val Ser Cys Ser Pro Arg Ser Pro
1 5 10 15

Glu Lys Ile Gln Pro Glu Leu Lys Val Leu Ala Asn Phe Glu Gly Ser
20 25 30

Tyr Trp Lys Gly Val Lys Gly Tyr Lys Ala Gln Glu Ala Phe Ala Lys
35 40 45

Glu Leu Ala Ala Leu Pro Gln Phe Leu Gly Thr Thr Tyr Lys Lys Glu
50 55 60

Ala Ala Phe Ser Thr Arg Asp Arg Val Ala Pro Met Lys Thr Tyr Gly
65 70 75 80

Phe Val Phe Val Asp Glu Glu Gly Tyr Leu Arg Ile Thr Glu Ala Gly
85 90 95

Lys Met Leu Ala Asn Asn Arg Arg Pro Lys Asp Val Phe Leu Lys Gln
100 105 110

Leu Val Lys Trp Gln Tyr Pro Ser Phe Gln His Lys Gly Lys Glu Tyr
115 120 125

Pro Glu Glu Glu Trp Ser Ile Asn Pro Leu Val Phe Val Leu Ser Leu
130 135 140

Leu Lys Lys Val Gly Gly Leu Ser Lys Leu Asp Ile Ala Met Phe Cys
145 150 155 160

Leu Thr Ala Thr Asn Asn Asn Gln Val Asp Glu Ile Ala Glu Glu Ile
165 170 175

Met Gln Phe Arg Asn Glu Arg Glu Lys Ile Lys Gly Gln Asn Lys Lys
180 185 190

```
Leu Glu Phe Thr Glu Asn Tyr Phe Phe Lys Arg Phe Glu Lys Ile Tyr
        195                 200                 205

Gly Asn Val Gly Lys Ile Arg Glu Gly Lys Ser Asp Ser Ser His Lys
    210                 215                 220

Ser Lys Ile Glu Thr Lys Met Arg Asn Ala Arg Asp Val Ala Asp Ala
225                 230                 235                 240

Thr Thr Arg Tyr Phe Arg Tyr Thr Gly Leu Phe Val Ala Arg Gly Asn
                245                 250                 255

Gln Leu Val Leu Asn Pro Glu Lys Ser Asp Leu Ile Asp Glu Ile Ile
            260                 265                 270

Ser Ser Ser Lys Val Val Lys Asn Tyr Thr Arg Val Glu Glu Phe His
        275                 280                 285

Glu Tyr Tyr Gly Asn Pro Ser Leu Pro Gln Phe Ser Phe Glu Thr Lys
    290                 295                 300

Glu Gln Leu Leu Asp Leu Ala His Arg Ile Arg Asp Glu Asn Thr Arg
305                 310                 315                 320

Leu Ala Glu Gln Leu Val Glu His Phe Pro Asn Val Lys Val Glu Ile
                325                 330                 335

Gln Val Leu Glu Asp Ile Tyr Asn Ser Leu Asn Lys Lys Val Asp Val
            340                 345                 350

Glu Thr Leu Lys Asp Val Ile Tyr His Ala Lys Glu Leu Gln Leu Glu
        355                 360                 365

Leu Lys Lys Lys Lys Leu Gln Ala Asp Phe Asn Asp Pro Arg Gln Leu
    370                 375                 380

Glu Glu Val Ile Asp Leu Leu Glu Val Tyr His Glu Lys Lys Asn Val
385                 390                 395                 400

Ile Glu Glu Lys Ile Lys Ala Arg Phe Ile Ala Asn Lys Asn Thr Val
                405                 410                 415

Phe Glu Trp Leu Thr Trp Asn Gly Phe Ile Ile Leu Gly Asn Ala Leu
            420                 425                 430

Glu Tyr Lys Asn Asn Phe Val Ile Asp Glu Glu Leu Gln Pro Val Thr
        435                 440                 445

His Ala Ala Gly Asn Gln Pro Asp Met Glu Ile Ile Tyr Glu Asp Phe
    450                 455                 460
```

Ile Val Leu Gly Glu Val Thr Thr Ser Lys Gly Ala Thr Gln Phe Lys
465 470 475 480

Met Glu Ser Glu Pro Val Thr Arg His Tyr Leu Asn Lys Lys Lys Glu
485 490 495

Leu Glu Lys Gln Gly Val Glu Lys Glu Leu Tyr Cys Leu Phe Ile Ala
500 505 510

Pro Glu Ile Asn Lys Asn Thr Phe Glu Glu Phe Met Lys Tyr Asn Ile
515 520 525

Val Gln Asn Thr Arg Ile Ile Pro Leu Ser Leu Lys Gln Phe Asn Met
530 535 540

Leu Leu Met Val Gln Lys Lys Leu Ile Glu Lys Gly Arg Arg Leu Ser
545 550 555 560

Ser Tyr Asp Ile Lys Asn Leu Met Val Ser Leu Tyr Arg Thr Thr Ile
565 570 575

Glu Cys Glu Arg Lys Tyr Thr Gln Ile Lys Ala Gly Leu Glu Glu Thr
580 585 590

Leu Asn Asn Trp Val Val Asp Lys Glu Val Arg Phe
595 600

<210> 39

<211> 186
<212> PRT
<213> Bacillus sp. D6

<220>
<223> ss.BspD6I (R.BspD6I small subunit)

<400> 39
Met Gln Asp Ile Leu Asp Phe Tyr Glu Glu Val Glu Lys Thr Ile Asn
1 5 10 15

Pro Pro Asn Tyr Phe Glu Trp Asn Thr Tyr Arg Val Phe Lys Lys Leu
20 25 30

Gly Ser Tyr Lys Asn Leu Val Pro Asn Phe Lys Leu Asp Asp Ser Gly
35 40 45

His Pro Ile Gly Asn Ala Ile Pro Gly Val Glu Asp Ile Leu Val Glu
50 55 60

Tyr Glu His Phe Ser Ile Leu Ile Glu Cys Ser Leu Thr Ile Gly Glu
65 70 75 80

Lys Gln Leu Asp Tyr Glu Gly Asp Ser Val Val Arg His Leu Gln Glu
85 90 95

Tyr Lys Lys Lys Gly Ile Glu Ala Tyr Thr Leu Phe Leu Gly Lys Ser
100 105 110

Ile Asp Leu Ser Phe Ala Arg His Ile Gly Phe Asn Lys Glu Ser Glu
115 120 125

Pro Val Ile Pro Leu Thr Val Asp Gln Phe Lys Lys Leu Val Thr Gln
130 135 140

Leu Lys Gly Asp Gly Glu His Phe Asn Pro Asn Lys Leu Lys Glu Ile
145 150 155 160

Leu Ile Lys Leu Leu Arg Ser Asp Leu Gly Tyr Asp Gln Ala Glu Glu
165 170 175

Trp Leu Thr Phe Ile Glu Tyr Asn Leu Lys
180 185

<210> 40

<211> 555
<212> PRT
<213> Paucimonas lemoignei

<220>
<223> R.PleI AAK27215.1

<400> 40
Met Ala Lys Pro Ile Asp Ser Lys Val Leu Phe Ile Thr Thr Ser Pro
1 5 10 15

Arg Thr Pro Glu Lys Met Val Pro Glu Ile Glu Leu Leu Asp Lys Asn
20 25 30

Phe Asn Gly Asp Val Trp Asn Lys Asp Thr Gln Thr Ala Phe Met Lys
35 40 45

Ile Leu Lys Glu Glu Ser Phe Phe Asp Gly Glu Gly Lys Asn Asp Pro
50 55 60

Ala Phe Ser Ala Arg Asp Arg Ile Asn Arg Ala Pro Lys Ser Leu Gly
65 70 75 80

Phe Val Ile Leu Thr Pro Lys Leu Ser Leu Thr Asp Ala Gly Val Glu
85 90 95

Leu Ile Lys Ala Lys Arg Lys Asp Asp Ile Phe Leu Arg Gln Met Leu
100 105 110

Lys Phe Gln Leu Pro Ser Pro Tyr His Lys Leu Ser Asp Lys Ala Ala
115 120 125

```
Leu Phe Tyr Val Lys Pro Tyr Leu Glu Ile Phe Arg Leu Val Arg His
    130                 135                 140

Phe Gly Ser Leu Thr Phe Asp Glu Leu Met Ile Phe Gly Leu Gln Ile
145                 150                 155                 160

Ile Asp Phe Arg Ile Phe Asn Gln Ile Val Asp Lys Ile Glu Asp Phe
                165                 170                 175

Arg Val Gly Lys Ile Glu Asn Lys Gly Arg Tyr Lys Thr Tyr Lys Lys
            180                 185                 190

Glu Arg Phe Glu Glu Glu Leu Gly Lys Ile Tyr Lys Asp Glu Leu Phe
        195                 200                 205

Gly Leu Thr Glu Ala Ser Ala Lys Thr Leu Ile Thr Lys Lys Gly Asn
    210                 215                 220

Asn Met Arg Asp Tyr Ala Asp Ala Cys Val Arg Tyr Leu Arg Ala Thr
225                 230                 235                 240

Gly Met Val Asn Val Ser Tyr Gln Gly Lys Ser Leu Ser Ile Val Gln
                245                 250                 255

Glu Lys Lys Glu Glu Val Asp Phe Phe Leu Lys Asn Thr Glu Arg Glu
            260                 265                 270

Pro Cys Phe Ile Asn Asp Glu Ala Ser Tyr Val Ser Tyr Leu Gly Asn
        275                 280                 285

Pro Asn Tyr Pro Lys Leu Phe Val Asp Asp Val Asp Arg Ile Lys Lys
    290                 295                 300

Lys Leu Arg Phe Asp Phe Lys Lys Thr Asn Lys Val Asn Ala Leu Thr
305                 310                 315                 320

Leu Pro Glu Leu Lys Glu Glu Leu Glu Asn Glu Ile Leu Ser Arg Lys
                325                 330                 335

Glu Asn Ile Leu Lys Ser Gln Ile Ser Asp Ile Lys Asn Phe Lys Leu
            340                 345                 350

Tyr Glu Asp Ile Gln Glu Val Phe Glu Lys Ile Glu Asn Asp Arg Thr
        355                 360                 365

Leu Ser Asp Ala Pro Leu Met Leu Glu Trp Asn Thr Trp Arg Ala Met
    370                 375                 380

Thr Met Leu Asp Gly Gly Glu Ile Lys Ala Asn Leu Lys Phe Asp Asp
385                 390                 395                 400
```

Phe Gly Ser Pro Met Ser Thr Ala Ile Gly Asn Met Pro Asp Ile Val
405 410 415

Cys Glu Tyr Asp Asp Phe Gln Leu Ser Val Glu Val Thr Met Ala Ser
420 425 430

Gly Gln Lys Gln Tyr Glu Met Glu Gly Glu Pro Val Ser Arg His Leu
435 440 445

Gly Lys Leu Lys Lys Ser Ser Glu Lys Pro Val Tyr Cys Leu Phe Ile
450 455 460

Ala Pro Lys Ile Asn Pro Ser Ser Val Ala His Phe Phe Met Ser His
465 470 475 480

Lys Val Asp Ile Glu Tyr Tyr Gly Gly Lys Ser Leu Ile Ile Pro Leu
485 490 495

Glu Leu Ser Val Phe Arg Lys Met Ile Glu Asp Thr Phe Lys Ala Ser
500 505 510

Tyr Ile Pro Lys Ser Asp Asn Val His Lys Leu Phe Lys Asn Phe Ala
515 520 525

Ser Ile Ala Asp Glu Ala Gly Asn Glu Lys Val Trp Tyr Glu Gly Val
530 535 540

Lys Arg Thr Ala Met Asn Trp Leu Ser Leu Ser
545 550 555

<210> 41

<211> 556
<212> PRT
<213> Micrococcus lylae

<220>
<223> MlyI AAK39546.1

<400> 41
Met Ala Ser Leu Ser Lys Thr Lys His Leu Phe Gly Phe Thr Ser Pro
1 5 10 15

Arg Thr Ile Glu Lys Ile Ile Pro Glu Leu Asp Ile Leu Ser Gln Gln
20 25 30

Phe Ser Gly Lys Val Trp Gly Glu Asn Gln Ile Asn Phe Phe Asp Ala
35 40 45

Ile Phe Asn Ser Asp Phe Tyr Glu Gly Thr Thr Tyr Pro Gln Asp Pro
50 55 60

```
Ala Leu Ala Ala Arg Asp Arg Ile Thr Arg Ala Pro Lys Ala Leu Gly
65                  70                  75                  80

Phe Ile Gln Leu Lys Pro Val Ile Gln Leu Thr Lys Ala Gly Asn Gln
                85                  90                  95

Leu Val Asn Gln Lys Arg Leu Pro Glu Leu Phe Thr Lys Gln Leu Leu
            100                 105                 110

Lys Phe Gln Leu Pro Ser Pro Tyr His Thr Gln Ser Pro Thr Val Asn
        115                 120                 125

Phe Asn Val Arg Pro Tyr Leu Glu Leu Leu Arg Leu Ile Asn Glu Leu
    130                 135                 140

Gly Ser Ile Ser Lys Thr Glu Ile Ala Leu Phe Phe Leu Gln Leu Val
145                 150                 155                 160

Asn Tyr Asn Lys Phe Asp Glu Ile Lys Asn Lys Ile Leu Lys Phe Arg
                165                 170                 175

Glu Thr Arg Lys Asn Asn Arg Ser Val Ser Trp Lys Thr Tyr Val Ser
            180                 185                 190

Gln Glu Phe Glu Lys Gln Ile Ser Ile Ile Phe Ala Asp Glu Val Thr
        195                 200                 205

Ala Lys Asn Phe Arg Thr Arg Glu Ser Ser Asp Glu Ser Phe Lys Lys
    210                 215                 220

Phe Val Lys Thr Lys Glu Gly Asn Met Lys Asp Tyr Ala Asp Ala Phe
225                 230                 235                 240

Phe Arg Tyr Ile Arg Gly Thr Gln Leu Val Thr Ile Asp Lys Asn Leu
                245                 250                 255

His Leu Lys Ile Ser Ser Leu Lys Gln Asp Ser Val Asp Phe Leu Leu
            260                 265                 270

Lys Asn Thr Asp Arg Asn Ala Leu Asn Leu Ser Leu Met Glu Tyr Glu
        275                 280                 285

Asn Tyr Leu Phe Asp Pro Asp Gln Leu Ile Val Leu Glu Asp Asn Ser
    290                 295                 300

Gly Leu Ile Asn Ser Lys Ile Lys Gln Leu Asp Asp Ser Ile Asn Val
305                 310                 315                 320

Glu Ser Leu Lys Ile Asp Asp Ala Lys Asp Leu Leu Asn Asp Leu Glu
                325                 330                 335
```

Ile Gln Arg Lys Ala Lys Thr Ile Glu Asp Thr Val Asn His Leu Lys
340 345 350

Leu Arg Ser Asp Ile Glu Asp Ile Leu Asp Val Phe Ala Lys Ile Lys
355 360 365

Lys Arg Asp Val Pro Asp Val Pro Leu Phe Leu Glu Trp Asn Ile Trp
370 375 380

Arg Ala Phe Ala Ala Leu Asn His Thr Gln Ala Ile Glu Gly Asn Phe
385 390 395 400

Ile Val Asp Leu Asp Gly Met Pro Leu Asn Thr Ala Pro Gly Lys Lys
405 410 415

Pro Asp Ile Glu Ile Asn Tyr Gly Ser Phe Ser Cys Ile Val Glu Val
420 425 430

Thr Met Ser Ser Gly Glu Thr Gln Phe Asn Met Glu Gly Ser Ser Val
435 440 445

Pro Arg His Tyr Gly Asp Leu Val Arg Lys Val Asp His Asp Ala Tyr
450 455 460

Cys Ile Phe Ile Ala Pro Lys Val Ala Pro Gly Thr Lys Ala His Phe
465 470 475 480

Phe Asn Leu Asn Arg Leu Ser Thr Lys His Tyr Gly Gly Lys Thr Lys
485 490 495

Ile Ile Pro Met Ser Leu Asp Asp Phe Ile Cys Phe Leu Gln Val Gly
500 505 510

Ile Thr His Asn Phe Gln Asp Ile Asn Lys Leu Lys Asn Trp Leu Asp
515 520 525

Asn Leu Ile Asn Phe Asn Leu Glu Ser Glu Asp Glu Glu Ile Trp Phe
530 535 540

Glu Glu Ile Ile Ser Lys Ile Ser Thr Trp Ala Ile
545 550 555

<210> 42

<211> 543
<212> PRT
<213> Geobacillus sp. Y412MC52

<220>
<223> AlwI YP_004134094.1

<400> 42
Met Asn Lys Lys Asn Thr Arg Lys Val Trp Phe Ile Thr Arg Pro Glu
1 5 10 15

```
Arg Asp Pro Arg Phe His Gln Glu Ala Leu Leu Ala Leu Gln Lys Ala
            20                  25                  30

Thr Asp Asp Phe Arg Leu Lys Trp Ala Gly Asn Arg Glu Val His Lys
        35                  40                  45

Arg Tyr Glu Glu Glu Leu Ala Asn Met Gly Ile Lys Arg Asn Asn Val
    50                  55                  60

Ser His Asp Gly Ser Gly Gly Arg Thr Trp Met Ala Met Leu Lys Thr
65                  70                  75                  80

Phe Ser Tyr Cys Tyr Val Asp Asp Asp Gly Tyr Ile Arg Leu Thr Lys
                85                  90                  95

Val Gly Glu Lys Leu Ile Gln Gly Glu Lys Val Tyr Glu Asn Thr Arg
            100                 105                 110

Lys Gln Val Leu Thr Leu Gln Tyr Pro Asn Ala Tyr Phe Leu Glu Pro
        115                 120                 125

Gly Phe Arg Pro Lys Phe Asp Glu Gly Phe Arg Ile Arg Pro Val Leu
    130                 135                 140

Phe Leu Ile Lys Leu Ala Asn Asp Glu Arg Leu Asp Phe Tyr Val Thr
145                 150                 155                 160

Lys Glu Glu Ile Thr Tyr Phe Ala Met Thr Ala Gln Lys Asp Ser Gln
                165                 170                 175

Leu Asp Glu Ile Val His Lys Ile Leu Ala Phe Arg Lys Ala Gly Pro
            180                 185                 190

Arg Glu Arg Glu Glu Met Lys Gln Asp Ile Ala Ala Lys Phe Asp His
        195                 200                 205

Arg Glu Arg Ser Asp Lys Gly Ala Arg Asp Phe Tyr Glu Ala His Ser
    210                 215                 220

Asp Val Ala His Thr Phe Met Leu Ile Ser Asp Tyr Thr Gly Leu Val
225                 230                 235                 240

Glu Tyr Ile Arg Gly Lys Ala Leu Lys Gly Asp Ser Ser Lys Ile Asn
                245                 250                 255

Glu Ile Lys Gln Glu Ile Ala Glu Ile Glu Lys Arg Tyr Pro Phe Asn
            260                 265                 270

Thr Arg Tyr Met Ile Ser Leu Glu Arg Met Ala Glu Asn Ser Gly Leu
        275                 280                 285
```

```
Asp Val Asp Ser Tyr Lys Ala Ser Arg Tyr Gly Asn Ile Lys Pro Ala
    290                 295                 300

Ala Asn Ser Ser Lys Leu Arg Ala Lys Ala Glu Arg Ile Leu Ala Gln
305                 310                 315                 320

Phe Pro Ser Ile Glu Ser Met Ser Lys Glu Glu Ile Ala Gly Ala Leu
                325                 330                 335

Gln Lys Tyr Leu Ser Pro Arg Asp Ile Glu Lys Val Ile His Glu Ile
            340                 345                 350

Val Glu Asn Lys Asp Asp Phe Glu Gly Ile Asn Ser Asp Phe Val Glu
        355                 360                 365

Thr Tyr Leu Asn Glu Lys Asp Asn Leu Ala Phe Glu Asp Lys Thr Gly
    370                 375                 380

Gln Ile Phe Ser Ala Leu Gly Phe Asp Val Ala Met Arg Pro Lys Ala
385                 390                 395                 400

Lys Asn Gly Glu Arg Thr Glu Ile Glu Ile Ile Ala Arg Tyr Gly Gly
                405                 410                 415

Ser Lys Phe Gly Ile Ile Asp Ala Lys Asn Tyr Ala Gly Lys Phe Pro
            420                 425                 430

Leu Ser Ser Ser Leu Val Ser His Met Ala Ser Glu Tyr Ile Pro Asn
        435                 440                 445

Tyr Thr Gly Tyr Glu Gly Lys Glu Leu Thr Phe Phe Gly Tyr Val Thr
    450                 455                 460

Ala Asn Asp Phe Ser Gly Glu Arg Asn Leu Glu Lys Ile Ser Asp Lys
465                 470                 475                 480

Ala Lys Arg Ile Thr Gly Asn Pro Ile Ser Gly Phe Leu Val Thr Ala
                485                 490                 495

Arg Thr Leu Leu Gly Phe Leu Asp Tyr Cys Ile Glu Asn Asp Val Pro
            500                 505                 510

Leu Glu Asp Arg Ala Glu Leu Phe Val Lys Ala Val Lys Asn Lys Gly
        515                 520                 525

Tyr Lys Ser Leu Glu Ala Leu Leu Arg Glu Leu Lys Glu Thr Ile
    530                 535                 540
```

<210> 43

<211> 685
<212> PRT
<213> Kocuria varians

<220>
<223> Mva1269I AAY97906.1

<400> 43

```
Met Tyr Leu Asn Thr Ala Val Phe Asn Ile Tyr Gly Asp Asn Ile Val
1               5                  10                  15

Glu Cys Ser Arg Ala Phe His Tyr Ile Leu Glu Gly Phe Lys Leu Ala
            20                  25                  30

Asn Ile Ser Ile Thr Gln Glu Tyr Asp Leu Gln Asn Ile Thr Thr Pro
        35                  40                  45

Lys Phe Cys Ile Tyr Thr Asp Lys Phe Arg Tyr Ile Phe Ile Phe Ile
    50                  55                  60

Pro Gly Thr Ser Ala Ser Arg Trp Asn Lys Asp Ile Tyr Lys Glu Leu
65                  70                  75                  80

Val Leu Asn Asn Gly Gly Pro Leu Lys Glu Gly Ala Asp Ala Ile Ile
                85                  90                  95

Thr Arg Ile Phe Ser Glu Asp Ser Glu Leu Val Leu Ala Ser Met Glu
            100                 105                 110

Phe Ser Ala Ala Leu Pro Ala Gly Asn Asn Thr Trp Gln Arg Ser Gly
        115                 120                 125

Arg Ala Tyr Ser Leu Thr Ala Ala Asn Ile Pro Tyr Phe Tyr Ile Val
    130                 135                 140

Gln Leu Gly Gly Lys Glu Ile Lys Lys Gly Lys Asp Gly Lys Ser Asp
145                 150                 155                 160

Lys Phe Ala Thr Arg Leu Pro Asn Pro Ala Leu Ser Leu Ser Phe Thr
                165                 170                 175

Leu Asn Thr Ile Lys Lys Pro Ala Pro Ser Leu Ile Val Tyr Asp Gln
            180                 185                 190

Ala Pro Glu Ala Asp Ser Ala Ile Ser Asp Leu Tyr Ser Asn Cys Tyr
        195                 200                 205

Gly Ile Asp Asp Phe Ser Leu Tyr Leu Phe Lys Leu Ile Thr Glu Glu
    210                 215                 220

Asn Asn Leu His Glu Leu Lys Asn Ile Tyr Asn Lys Asn Val Glu Phe
225                 230                 235                 240
```

```
Leu Gln Leu Arg Ser Val Asp Glu Lys Gly Lys Asn Phe Ser Gly Lys
                245                 250                 255

Asp Tyr Lys Tyr Ile Phe Glu His Lys Asp Pro Tyr Lys Gly Leu Thr
            260                 265                 270

Glu Val Val Lys Glu Arg Lys Ile Pro Trp Lys Lys Lys Thr Ala Thr
        275                 280                 285

Lys Thr Phe Glu Asn Phe Pro Leu Arg Asn Gln Ala Pro Ile Phe Arg
    290                 295                 300

Leu Ile Asp Phe Leu Ser Thr Lys Ser Tyr Gly Ile Val Ser Lys Asp
305                 310                 315                 320

Ser Leu Pro Leu Thr Phe Ile Pro Ser Glu His Arg Val Glu Val Ala
                325                 330                 335

Asn Tyr Ile Cys Asn Gln Leu Tyr Ile Asp Lys Val Ser Asp Glu Phe
            340                 345                 350

Val Lys Trp Ile Tyr Lys Lys Glu Asp Leu Ala Ile Cys Ile Ile Asn
        355                 360                 365

Gly Phe Lys Pro Gly Gly Asp Asp Ser Arg Pro Asp Arg Gly Leu Pro
    370                 375                 380

Pro Phe Thr Lys Met Leu Thr Asn Leu Asp Ile Leu Thr Leu Met Phe
385                 390                 395                 400

Gly Pro Ala Pro Pro Thr Gln Trp Asp Tyr Leu Asp Ser Asp Pro Glu
                405                 410                 415

Lys Leu Asn Lys Thr Asn Gly Leu Trp Gln Ser Ile Phe Ala Phe Ser
            420                 425                 430

Asp Ala Ile Leu Val Asp Ser Ser Thr Arg Asp Asn Asn Lys Phe Val
        435                 440                 445

Tyr Asn Ala Tyr Leu Lys Glu His Trp Val Val Gln Arg Glu Lys Lys
    450                 455                 460

Glu Ser Asn Thr Pro Ile Ser Tyr Phe Pro Lys Ser Val Gly Glu His
465                 470                 475                 480

Asp Val Asp Thr Ser Leu His Ile Leu Phe Thr Tyr Ile Gly Lys His
                485                 490                 495

Phe Glu Ser Ala Cys Asn Pro Pro Gly Gly Asp Trp Ser Gly Val Ser
            500                 505                 510
```

Leu Leu Lys Asn Asn Ile Glu Tyr Arg Trp Thr Ser Met Tyr Arg Val
515 520 525

Ser Gln Asp Gly Thr Lys Arg Pro Asp His Ile Tyr Gln Leu Val Tyr
530 535 540

Asn Ser Thr Asp Thr Leu Leu Leu Ile Glu Ser Lys Gly Ile Lys Asn
545 550 555 560

Asp Leu Leu Lys Ser Lys Glu Ala Asn Val Gly Ile Gly Met Ile Asn
565 570 575

Tyr Leu Lys Asn Leu Met Ala Arg Asp Tyr Thr Ala Val Lys Lys Asp
580 585 590

Gly Glu Trp Lys Asn Ile His Gly Gln Met Thr Leu Asp Lys Phe Leu
595 600 605

Thr Phe Ser Ala Val Ala Tyr Leu Phe Thr Thr Asp Phe Asp Asn Glu
610 615 620

Tyr Thr Ser Ala Ala Glu Leu Leu Val His Ser Asn Thr Gln Leu Ala
625 630 635 640

Phe Ala Leu Glu Ile Lys Glu Lys Asn Ser Val Met His Ile Phe Thr
645 650 655

Ala Asn Thr Val Ala Tyr Asn Phe Ala Glu Tyr Leu Leu Glu Thr Met
660 665 670

Arg Asn Ser His Leu Pro Leu Lys Ile Tyr Lys Pro Ile
675 680 685

<210> 44

<211> 599
<212> PRT
<213> Geobacillus stearothermophilus

<220>
<223> BsrI ADR72996.1

<400> 44
Met Arg Asn Ile Arg Ile Tyr Ser Glu Val Lys Glu Gln Gly Ile Phe
1 5 10 15

Phe Lys Glu Val Ile Gln Ser Val Leu Glu Lys Ala Asn Val Glu Val
20 25 30

Val Leu Val Asn Ser Ala Met Leu Asp Tyr Ser Asp Val Ser Val Ile
35 40 45

Ser Leu Ile Arg Asn Gln Lys Lys Phe Asp Leu Leu Val Ser Glu Val
50 55 60

```
Arg Asp Lys Arg Glu Ile Pro Ile Val Met Val Glu Phe Ser Thr Ala
65                  70                  75                  80

Val Thr Thr Asp Asp His Glu Leu Gln Arg Ala Asp Ala Met Phe Trp
                85                  90                  95

Ala Tyr Lys Tyr Lys Ile Pro Tyr Leu Lys Ile Ser Pro Met Glu Lys
            100                 105                 110

Lys Ser Gln Thr Ala Asp Asp Lys Phe Gly Gly Gly Arg Leu Leu Ser
        115                 120                 125

Val Asn Asp Gln Ile Ile His Met Tyr Arg Thr Asp Gly Val Met Tyr
    130                 135                 140

His Ile Glu Trp Glu Ser Met Asp Asn Ser Ala Tyr Val Lys Asn Ala
145                 150                 155                 160

Glu Leu Tyr Pro Ser Cys Pro Asp Cys Ala Pro Glu Leu Ala Ser Leu
                165                 170                 175

Phe Arg Cys Leu Leu Glu Thr Ile Glu Lys Cys Glu Asn Ile Glu Asp
            180                 185                 190

Tyr Tyr Arg Ile Leu Leu Asp Lys Leu Gly Lys Gln Lys Val Ala Val
        195                 200                 205

Lys Trp Gly Asn Phe Arg Glu Glu Lys Thr Leu Glu Gln Trp Lys His
    210                 215                 220

Glu Lys Phe Asp Leu Leu Glu Arg Phe Ser Lys Ser Ser Ser Arg Met
225                 230                 235                 240

Glu Tyr Asp Lys Asp Lys Lys Glu Leu Lys Ile Lys Val Asn Arg Tyr
                245                 250                 255

Gly His Ala Met Asp Pro Glu Arg Gly Ile Leu Ala Phe Trp Lys Leu
            260                 265                 270

Val Leu Gly Asp Glu Trp Lys Ile Val Ala Glu Phe Gln Leu Gln Arg
        275                 280                 285

Lys Thr Leu Lys Gly Arg Gln Ser Tyr Gln Ser Leu Phe Asp Glu Val
    290                 295                 300

Ser Gln Glu Glu Lys Leu Met Asn Ile Ala Ser Glu Ile Ile Lys Asn
305                 310                 315                 320

Gly Asn Val Ile Ser Pro Asp Lys Ala Ile Glu Ile His Lys Leu Ala
                325                 330                 335
```

```
Thr Ser Ser Thr Met Ile Ser Thr Ile Asp Leu Gly Thr Pro Glu Arg
            340                 345                 350

Lys Tyr Ile Thr Asp Asp Ser Leu Lys Gly Tyr Leu Gln His Gly Leu
        355                 360                 365

Ile Thr Asn Ile Tyr Lys Asn Leu Leu Tyr Tyr Val Asp Glu Ile Arg
    370                 375                 380

Phe Thr Asp Leu Gln Arg Lys Thr Ile Ala Ser Leu Thr Trp Asn Lys
385                 390                 395                 400

Glu Ile Val Asn Asp Tyr Tyr Lys Ser Leu Met Asp Gln Leu Leu Asp
                405                 410                 415

Lys Asn Leu Arg Val Leu Pro Leu Thr Ser Ile Lys Asn Ile Ser Glu
            420                 425                 430

Asp Leu Ile Thr Trp Ser Ser Lys Glu Ile Leu Ile Asn Leu Gly Tyr
        435                 440                 445

Lys Ile Leu Ala Ala Ser Tyr Pro Glu Ala Gln Gly Asp Arg Cys Ile
    450                 455                 460

Leu Val Gly Pro Thr Gly Lys Lys Thr Glu Arg Lys Phe Ile Asp Leu
465                 470                 475                 480

Ile Ala Ile Ser Pro Lys Ser Lys Gly Val Ile Leu Leu Glu Cys Lys
                485                 490                 495

Asp Lys Leu Ser Lys Ser Lys Asp Asp Cys Glu Lys Met Asn Asp Leu
            500                 505                 510

Leu Asn His Asn Tyr Asp Lys Val Thr Lys Leu Ile Asn Val Leu Asn
        515                 520                 525

Ile Asn Asn Tyr Asn Tyr Asn Asn Ile Ile Tyr Thr Gly Val Ala Gly
    530                 535                 540

Leu Ile Gly Arg Lys Asn Val Asp Asn Leu Pro Val Asp Phe Val Ile
545                 550                 555                 560

Lys Phe Lys Tyr Asp Ala Lys Asn Leu Lys Leu Asn Trp Glu Ile Asn
                565                 570                 575

Ser Asp Ile Leu Gly Lys His Ser Gly Ser Phe Ser Met Glu Asp Val
            580                 585                 590

Ala Val Val Arg Lys Arg Ser
        595
```

<210> 45

<211> 676
<212> PRT
<213> Geobacillus stearothermophilus

<220>
<223> BsmI AAL86024.1

<400> 45
Met Asn Val Phe Arg Ile His Gly Asp Asn Ile Ile Glu Cys Glu Arg
1 5 10 15

Val Ile Asp Leu Ile Leu Ser Lys Ile Asn Pro Gln Lys Val Lys Arg
20 25 30

Gly Phe Ile Ser Leu Ser Cys Pro Phe Ile Glu Ile Ile Phe Lys Glu
35 40 45

Gly His Asp Tyr Phe His Trp Arg Phe Asp Met Phe Pro Gly Phe Asn
50 55 60

Lys Asn Thr Asn Asp Arg Trp Asn Ser Asn Ile Leu Asp Leu Leu Ser
65 70 75 80

Gln Lys Gly Ser Phe Leu Tyr Glu Thr Pro Asp Val Ile Ile Thr Ser
85 90 95

Leu Asn Asn Gly Lys Glu Glu Ile Leu Met Ala Ile Glu Phe Cys Ser
100 105 110

Ala Leu Gln Ala Gly Asn Gln Ala Trp Gln Arg Ser Gly Arg Ala Tyr
115 120 125

Ser Val Gly Arg Thr Gly Tyr Pro Tyr Ile Tyr Ile Val Asp Phe Val
130 135 140

Lys Tyr Glu Leu Asn Asn Ser Asp Arg Ser Arg Lys Asn Leu Arg Phe
145 150 155 160

Pro Asn Pro Ala Ile Pro Tyr Ser Tyr Ile Ser His Ser Lys Asn Thr
165 170 175

Gly Asn Phe Ile Val Gln Ala Tyr Phe Arg Gly Glu Glu Tyr Gln Pro
180 185 190

Lys Tyr Asp Lys Lys Leu Lys Phe Phe Asp Glu Thr Ile Phe Ala Glu
195 200 205

Asp Asp Ile Ala Asp Tyr Ile Ile Ala Lys Leu Gln His Arg Asp Thr
210 215 220

```
Ser Asn Ile Glu Gln Leu Leu Ile Asn Lys Asn Leu Lys Met Val Glu
225                 230                 235                 240

Phe Leu Ser Lys Asn Thr Lys Asn Asp Asn Asn Phe Thr Tyr Ser Glu
                245                 250                 255

Trp Glu Ser Ile Tyr Asn Gly Thr Tyr Arg Ile Thr Asn Leu Pro Ser
            260                 265                 270

Leu Gly Arg Phe Lys Phe Arg Lys Lys Ile Ala Glu Lys Ser Leu Ser
        275                 280                 285

Gly Lys Val Lys Glu Phe Asn Asn Ile Val Gln Arg Tyr Ser Val Gly
    290                 295                 300

Leu Ala Ser Ser Asp Leu Pro Phe Gly Val Ile Arg Lys Glu Ser Arg
305                 310                 315                 320

Asn Asp Phe Ile Asn Asp Val Cys Lys Leu Tyr Asn Ile Asn Asp Met
                325                 330                 335

Lys Ile Ile Lys Glu Leu Lys Glu Asp Ala Asp Leu Ile Val Cys Met
            340                 345                 350

Leu Lys Gly Phe Lys Pro Arg Gly Asp Asp Asn Arg Pro Asp Arg Gly
        355                 360                 365

Ala Leu Pro Leu Val Ala Met Leu Ala Gly Glu Asn Ala Gln Ile Phe
    370                 375                 380

Thr Phe Ile Tyr Gly Pro Leu Ile Lys Gly Ala Ile Asn Leu Ile Asp
385                 390                 395                 400

Gln Asp Ile Asn Lys Leu Ala Lys Arg Asn Gly Leu Trp Lys Ser Phe
                405                 410                 415

Val Ser Leu Ser Asp Phe Ile Val Leu Asp Cys Pro Ile Ile Gly Glu
            420                 425                 430

Ser Tyr Asn Glu Phe Arg Leu Ile Ile Asn Lys Asn Asn Lys Glu Ser
        435                 440                 445

Ile Leu Arg Lys Thr Ser Lys Gln Gln Asn Ile Leu Val Asp Pro Thr
    450                 455                 460

Pro Asn His Tyr Gln Glu Asn Asp Val Asp Thr Val Ile Tyr Ser Ile
465                 470                 475                 480

Phe Lys Tyr Ile Val Pro Asn Cys Phe Ser Gly Met Cys Asn Pro Pro
                485                 490                 495
```

Gly Gly Asp Trp Ser Gly Leu Ser Ile Ile Arg Asn Gly His Glu Phe
500 505 510

Arg Trp Leu Ser Leu Pro Arg Val Ser Glu Asn Gly Lys Arg Pro Asp
515 520 525

His Val Ile Gln Ile Leu Asp Leu Phe Glu Lys Pro Leu Leu Leu Ser
530 535 540

Ile Glu Ser Lys Glu Lys Pro Asn Asp Leu Glu Pro Lys Ile Gly Val
545 550 555 560

Gln Leu Ile Lys Tyr Ile Glu Tyr Leu Phe Asp Phe Thr Pro Ser Val
565 570 575

Gln Arg Lys Ile Ala Gly Gly Asn Trp Glu Phe Gly Asn Lys Ser Leu
580 585 590

Val Pro Asn Asp Phe Ile Leu Leu Ser Ala Gly Ala Phe Ile Asp Tyr
595 600 605

Asp Asn Leu Thr Glu Asn Asp Tyr Glu Lys Ile Phe Glu Val Thr Gly
610 615 620

Cys Asp Leu Leu Ile Ala Ile Lys Asn Gln Asn Asn Pro Gln Lys Trp
625 630 635 640

Val Ile Lys Phe Lys Pro Lys Asn Thr Ile Ala Glu Lys Leu Val Asn
645 650 655

Tyr Ile Lys Leu Asn Phe Lys Ser Asn Ile Phe Asp Thr Gly Phe Phe
660 665 670

His Ile Glu Gly
675

<210> 46

<211> 465
<212> PRT
<213> artificial sequence

<220>
<223> Nb.BtsCI ADI24225.1

<400> 46
Met Lys Arg Ile Leu Tyr Leu Leu Thr Glu Glu Arg Pro Lys Ile Asn
1 5 10 15

Ile Ile His Gln Ile Ile Asn Leu Glu Tyr Lys Ala Thr Leu His Phe
20 25 30

Gly Ala Lys Ile Val Pro Val Met Asn Glu Glu Asn Lys Phe Thr Phe
35 40 45

```
Ile Tyr His Val Lys Gly Ile Glu Val Glu Gly Phe Asp Ala Val Leu
    50                  55                  60

Ile Lys Ile Val Ser Gly His Ser Ser Phe Val Asp Tyr Leu Val Phe
65                  70                  75                  80

Asp Ser Asn Asp Leu Lys Pro Glu Lys Asn Thr Ile Thr Leu Phe Asp
                85                  90                  95

Leu Asp Gln Tyr Glu Leu Asp Leu Ser Tyr Tyr Phe Gly Lys Gly Trp
            100                 105                 110

Ile Val Arg Ile Pro Ser Pro Ser Asp Leu Pro Lys Tyr Val Val Glu
        115                 120                 125

Glu Thr Lys Thr Asp Asp His Glu Ser Arg Asn Thr Asn Ala Tyr Gln
    130                 135                 140

Arg Ser Ser Lys Phe Val Phe Cys Glu Leu Tyr Tyr Gly Lys Glu Val
145                 150                 155                 160

Lys Lys Tyr Met Leu Tyr Asp Ile Ser Asp Gly Arg Thr Leu Ser Gly
                165                 170                 175

Thr Asp Thr His Asn Phe Gly Met Arg Met Leu Val Thr Asn Asn Val
            180                 185                 190

Asn Leu Val Gly Val Pro Asn Met Tyr Leu Pro Phe Thr Asp Ile Lys
        195                 200                 205

Glu Phe Ile Asn Glu Lys Asn Arg Ile Ala Asp Asn Gly Pro Ser His
    210                 215                 220

Asn Val Pro Ile Arg Leu Lys Leu Asp Lys Glu Lys Asn Val Ile Tyr
225                 230                 235                 240

Ile Ser Ala Lys Leu Asp Lys Gly Asn Gly Lys Asn Lys Asn Lys Ile
                245                 250                 255

Ser Asn Asp Pro Asn Ile Gly Ala Val Ala Ile Ile Ser Ala Thr Leu
            260                 265                 270

Arg Asn Leu Asn Trp Lys Gly Asp Ile Glu Ile Ile Asn His Asn Leu
        275                 280                 285

Leu Pro Ser Ser Ile Ser Ser Arg Ser Asn Gly Asn Lys Leu Leu Tyr
    290                 295                 300

Ile Met Lys Lys Leu Gly Val Arg Phe Asn Asn Ile Asn Val Asn Trp
305                 310                 315                 320
```

Asn Asn Ile Lys Asn Asn Ile Asn Tyr Phe Phe Tyr Asn Ile Thr Ser
325 330 335

Glu Lys Ile Val Ser Ile Tyr Tyr His Leu Tyr Val Glu Asp Lys Leu
340 345 350

Ser Asn Ala Arg Val Ile Phe Asp Asn His Ala Gly Cys Gly Lys Ser
355 360 365

Tyr Phe Arg Thr Leu Asn Asn Lys Ile Ile Pro Val Gly Lys Glu Ile
370 375 380

Pro Leu Pro Ala Leu Val Ile Phe Asp Ser Asp Gln Asn Ile Val Lys
385 390 395 400

Val Ile Ala Ala Ala Lys Ala Glu Asn Val Tyr Asn Gly Val Glu Gln
405 410 415

Leu Ser Thr Phe Asp Lys Phe Ile Glu Ser Tyr Ile Asn Lys Tyr Tyr
420 425 430

Pro Gly Ala Ala Val Glu Cys Ser Val Ile Thr Trp Gly Lys Ser Ser
435 440 445

Asn Pro Tyr Val Ser Phe Tyr Leu Asp Lys Asp Gly Ser Ala Val Phe
450 455 460

Leu
465

<210> 47

<211> 465
<212> PRT
<213> artificial sequence

<220>
<223> Nt.BtsCI ADI24224.1

<400> 47
Met Lys Arg Ile Leu Tyr Leu Leu Thr Glu Glu Arg Pro Lys Ile Asn
1 5 10 15

Ile Ile His Gln Ile Ile Asn Leu Glu Tyr Lys Ala Thr Leu His Phe
20 25 30

Gly Ala Lys Ile Val Pro Val Met Asn Glu Glu Asn Lys Phe Thr Phe
35 40 45

Ile Tyr His Val Lys Gly Ile Glu Val Glu Gly Phe Asp Ala Val Leu
50 55 60

```
Ile Lys Ile Val Ser Gly His Ser Ser Phe Val Asp Tyr Leu Val Phe
65                  70                  75                  80

Asp Ser Asn Asp Leu Lys Pro Glu Lys Asn Thr Ile Thr Leu Phe Asp
                85                  90                  95

Leu Asp Gln Tyr Glu Leu Asp Leu Ser Tyr Tyr Phe Gly Lys Gly Trp
            100                 105                 110

Ile Val Arg Ile Pro Ser Pro Ser Asp Leu Pro Lys Tyr Val Val Phe
        115                 120                 125

Glu Thr Lys Thr Asp Asp His Glu Ser Arg Asn Thr Asn Ala Tyr Gln
    130                 135                 140

Arg Ser Ser Lys Phe Val Phe Cys Glu Leu Tyr Tyr Gly Lys Glu Val
145                 150                 155                 160

Lys Lys Tyr Met Leu Tyr Asp Ile Ser Asp Gly Arg Thr Leu Ser Gly
                165                 170                 175

Thr Asp Thr His Asn Phe Gly Met Arg Met Leu Val Thr Asn Asn Val
            180                 185                 190

Asn Leu Val Gly Val Pro Asn Met Tyr Leu Pro Phe Thr Asp Ile Lys
        195                 200                 205

Glu Phe Ile Asn Glu Lys Asn Arg Ile Ala Asp Asn Gly Pro Ser His
    210                 215                 220

Asn Val Pro Ile Arg Leu Lys Leu Asp Lys Glu Lys Asn Val Ile Tyr
225                 230                 235                 240

Ile Ser Ala Lys Leu Asp Lys Gly Asn Gly Lys Asn Lys Asn Lys Ile
                245                 250                 255

Ser Asn Asp Pro Asn Ile Gly Ala Val Ala Ile Ile Ser Ala Thr Leu
            260                 265                 270

Arg Asn Leu Asn Trp Lys Gly Asp Ile Glu Ile Ile Asn His Asn Leu
        275                 280                 285

Leu Pro Ser Ser Ile Ser Ser Arg Ser Asn Gly Asn Lys Leu Leu Tyr
    290                 295                 300

Ile Met Lys Lys Leu Gly Val Arg Phe Asn Asn Ile Asn Val Asn Trp
305                 310                 315                 320

Asn Asn Ile Lys Asn Asn Ile Asn Tyr Phe Phe Tyr Asn Ile Thr Ser
                325                 330                 335
```

Glu Lys Ile Val Ser Ile Tyr Tyr His Leu Tyr Val Glu Asp Lys Leu
340 345 350

Ser Asn Ala Arg Val Ile Phe Asp Asn His Ala Gly Cys Gly Lys Ser
355 360 365

Tyr Phe Arg Thr Leu Asn Asn Lys Ile Ile Pro Val Gly Lys Glu Ile
370 375 380

Pro Leu Pro Asp Leu Val Ile Phe Asp Ser Asp Gln Asn Ile Val Lys
385 390 395 400

Val Ile Glu Ala Glu Lys Ala Glu Asn Val Tyr Asn Gly Val Glu Gln
405 410 415

Leu Ser Thr Phe Asp Lys Phe Ile Glu Ser Tyr Ile Asn Lys Tyr Tyr
420 425 430

Pro Gly Ala Ala Val Glu Cys Ser Val Ile Thr Trp Gly Lys Ser Ser
435 440 445

Asn Pro Tyr Val Ser Phe Tyr Leu Asp Lys Asp Gly Ser Ala Val Phe
450 455 460

Leu
465

<210> 48

<211> 164
<212> PRT
<213> Geobacillus thermoglucosidasius

<220>
<223> R1.BtsI ABC75874.1

<400> 48
Met Lys Ile Thr Glu Gly Ile Val His Val Ala Met Arg His Phe Leu
1 5 10 15

Lys Ser Asn Gly Trp Lys Leu Ile Ala Gly Gln Tyr Pro Gly Gly Ser
20 25 30

Asp Asp Glu Leu Thr Ala Leu Asn Ile Val Asp Pro Val Val Ala Arg
35 40 45

Asp Asn Ser Pro Asp Pro Arg Arg His Ser Leu Gly Lys Ile Val Pro
50 55 60

Asp Leu Ile Ala Tyr Lys Asn Asp Asp Leu Leu Val Ile Glu Ala Lys
65 70 75 80

Pro Lys Tyr Ser Gln Asp Asp Arg Asp Lys Leu Leu Tyr Leu Leu Ser
85 90 95

Glu Arg Lys His Asp Phe Tyr Ala Ala Leu Glu Lys Phe Ala Thr Glu
100 105 110

Arg Asn His Pro Glu Leu Leu Pro Val Ser Lys Leu Asn Ile Ile Pro
115 120 125

Gly Leu Ala Phe Ser Ala Ser Glu Asn Lys Phe Lys Lys Asp Pro Gly
130 135 140

Phe Val Tyr Ile Arg Val Ser Gly Ile Phe Glu Ala Phe Met Glu Gly
145 150 155 160

Tyr Asp Trp Gly

<210> 49

<211> 328
<212> PRT
<213> Geobacillus thermoglucosidasius

<220>
<223> R2.BtsI ABC75876.1

<400> 49
Met Gln Ile Glu Gln Leu Met Lys Ser Leu Thr Ile Tyr Phe Asp Asp
1 5 10 15

Ile Gln Glu Gly Leu Trp Phe Lys Asn Leu His Pro Leu Leu Glu Ser
20 25 30

Ala Ser Leu Glu Ala Ile Thr Gly Ser Leu Lys Arg Asn Pro Asn Leu
35 40 45

Ala Asp Val Leu Lys Tyr Asp Arg Pro Asp Ile Ile Leu Thr Leu Asn
50 55 60

Gln Thr Pro Ile Leu Val Ile Glu Arg Thr Ile Glu Val Pro Ser Gly
65 70 75 80

His Asn Val Gly Gln Arg Tyr Gly Arg Leu Ala Ala Ala Ser Glu Ala
85 90 95

Gly Val Pro Leu Val Tyr Phe Gly Pro Tyr Ala Ala Arg Lys His Gly
100 105 110

Gly Ala Thr Glu Gly Pro Arg Tyr Met Asn Leu Arg Leu Phe Tyr Ala
115 120 125

Leu Asp Val Met Gln Lys Val Asn Gly Ser Ala Ile Thr Thr Ile Asn
130 135 140

Trp Pro Val Asp Gln Asn Phe Glu Ile Leu Gln Asp Pro Ser Lys Asp
145 150 155 160

Lys Arg Met Lys Glu Tyr Leu Glu Met Phe Phe Asp Asn Leu Leu Lys
165 170 175

Tyr Gly Ile Ala Gly Ile Asn Leu Ala Ile Arg Asn Ser Ser Phe Gln
180 185 190

Ala Glu Gln Leu Ala Glu Arg Glu Lys Phe Val Glu Thr Met Ile Thr
195 200 205

Asn Pro Glu Gln Tyr Asp Val Pro Pro Asp Ser Val Gln Ile Leu Asn
210 215 220

Ala Glu Arg Phe Phe Asn Glu Leu Gly Ile Ser Glu Asn Lys Arg Ile
225 230 235 240

Ile Cys Asp Glu Val Val Leu Tyr Gln Val Gly Met Thr Tyr Val Arg
245 250 255

Ser Asp Pro Tyr Thr Gly Met Ala Leu Leu Tyr Lys Tyr Leu Tyr Ile
260 265 270

Leu Gly Ser Glu Arg Asn Arg Cys Leu Ile Leu Lys Phe Pro Asn Ile
275 280 285

Thr Thr Asp Met Trp Lys Lys Val Ala Phe Gly Ser Arg Glu Arg Lys
290 295 300

Asp Val Arg Ile Tyr Arg Ser Val Ser Asp Gly Ile Leu Phe Ala Asp
305 310 315 320

Gly Tyr Leu Ser Lys Glu Glu Leu
325

<210> 50

<211> 275
<212> PRT
<213> Brevibacillus brevis

<220>
<223> BbvCI subunit 1 AAX14652.1

<400> 50
Met Ile Asn Glu Asp Phe Phe Ile Tyr Glu Gln Leu Ser His Lys Lys
1 5 10 15

Asn Leu Glu Gln Lys Gly Lys Asn Ala Phe Asp Glu Glu Thr Glu Glu
20 25 30

Leu Val Arg Gln Ala Lys Ser Gly Tyr His Ala Phe Ile Glu Gly Ile
35 40 45

```
Asn Tyr Asp Glu Val Thr Lys Leu Asp Leu Asn Ser Ser Val Ala Ala
    50                  55                  60

Leu Glu Asp Tyr Ile Ser Ile Ala Lys Glu Ile Glu Lys Lys His Lys
65                  70                  75                  80

Met Phe Asn Trp Arg Ser Asp Tyr Ala Gly Ser Ile Ile Pro Glu Phe
                85                  90                  95

Leu Tyr Arg Ile Val His Val Ala Thr Val Lys Ala Gly Leu Lys Pro
            100                 105                 110

Ile Phe Ser Thr Arg Asn Thr Ile Ile Glu Ile Ser Gly Ala Ala His
        115                 120                 125

Arg Glu Gly Leu Gln Ile Arg Arg Lys Asn Glu Asp Phe Ala Leu Gly
    130                 135                 140

Phe His Glu Val Asp Val Lys Ile Ala Ser Glu Ser His Arg Val Ile
145                 150                 155                 160

Ser Leu Ala Val Ala Cys Glu Val Lys Thr Asn Ile Asp Lys Asn Lys
                165                 170                 175

Leu Asn Gly Leu Asp Phe Ser Ala Glu Arg Met Lys Arg Thr Tyr Pro
            180                 185                 190

Gly Ser Ala Tyr Phe Leu Ile Thr Glu Thr Leu Asp Phe Ser Pro Asp
        195                 200                 205

Glu Asn His Ser Ser Gly Leu Ile Asp Glu Ile Tyr Val Leu Arg Lys
    210                 215                 220

Gln Val Arg Thr Lys Asn Arg Val Gln Lys Ala Pro Leu Cys Pro Ser
225                 230                 235                 240

Val Phe Ala Glu Leu Leu Glu Asp Ile Leu Glu Ile Ser Tyr Arg Ala
                245                 250                 255

Ser Asn Val Lys Gly His Val Tyr Asp Arg Leu Glu Gly Gly Lys Leu
            260                 265                 270

Ile Arg Val
        275
```

<210> 51

<211> 285
<212> PRT
<213> Brevibacillus brevis

<220>
<223> BbvCI subunit 2 AAX14653.1

<400> 51

```
Met Phe Asn Gln Phe Asn Pro Leu Val Tyr Thr His Gly Gly Lys Leu
1               5                  10                  15

Glu Arg Lys Ser Lys Lys Asp Lys Thr Ala Ser Lys Val Phe Glu Glu
            20                  25                  30

Phe Gly Val Met Glu Ala Tyr Asn Cys Trp Lys Glu Ala Ser Leu Cys
        35                  40                  45

Ile Gln Gln Arg Asp Lys Asp Ser Val Leu Lys Leu Val Ala Ala Leu
    50                  55                  60

Asn Thr Tyr Lys Asp Ala Val Glu Pro Ile Phe Asp Ser Arg Leu Asn
65                  70                  75                  80

Ser Ala Gln Glu Val Leu Gln Pro Ser Ile Leu Glu Glu Phe Phe Glu
                85                  90                  95

Tyr Leu Phe Ser Arg Ile Asp Ser Ile Val Gly Val Asn Ile Pro Ile
            100                 105                 110

Arg His Pro Ala Lys Gly Tyr Leu Ser Leu Ser Phe Asn Pro His Asn
        115                 120                 125

Ile Glu Thr Leu Ile Gln Ser Pro Glu Tyr Thr Val Arg Ala Lys Asp
    130                 135                 140

His Asp Phe Ile Ile Gly Gly Ser Ala Lys Leu Thr Ile Gln Gly His
145                 150                 155                 160

Gly Gly Glu Gly Glu Thr Thr Asn Ile Val Val Pro Ala Val Ala Ile
                165                 170                 175

Glu Cys Lys Arg Tyr Leu Glu Arg Asn Met Leu Asp Glu Cys Ala Gly
            180                 185                 190

Thr Ala Glu Arg Leu Lys Arg Ala Thr Pro Tyr Cys Leu Tyr Phe Val
        195                 200                 205

Val Ala Glu Tyr Leu Lys Leu Asp Asp Gly Ala Pro Glu Leu Thr Glu
    210                 215                 220

Ile Asp Glu Ile Tyr Ile Leu Arg His Gln Arg Asn Ser Glu Arg Asn
225                 230                 235                 240

Lys Pro Gly Phe Lys Pro Asn Pro Ile Asp Gly Glu Leu Ile Trp Asp
                245                 250                 255
```

Leu Tyr Gln Glu Val Met Asn His Leu Gly Lys Ile Trp Trp Asp Pro
260 265 270

Asn Ser Ala Leu Gln Arg Gly Lys Val Phe Asn Arg Pro
275 280 285

<210> 52

<211> 294
<212> PRT
<213> Bacillus pumilus

<220>
<223> Bpu10I alpha subunit CAA74998.1

<400> 52
Met Gly Val Glu Gln Glu Trp Ile Lys Asn Ile Thr Asp Met Tyr Gln
1 5 10 15

Ser Pro Glu Leu Ile Pro Ser His Ala Ser Asn Leu Leu His Gln Leu
20 25 30

Lys Arg Glu Lys Arg Asn Glu Lys Leu Lys Lys Ala Leu Glu Ile Ile
35 40 45

Thr Pro Asn Tyr Ile Ser Tyr Ile Ser Ile Leu Leu Asn Asn His Asn
50 55 60

Met Thr Arg Lys Glu Ile Val Ile Leu Val Asp Ala Leu Asn Glu Tyr
65 70 75 80

Met Asn Thr Leu Arg His Pro Ser Val Lys Ser Val Phe Ser His Gln
85 90 95

Ala Asp Phe Tyr Ser Ser Val Leu Pro Glu Phe Phe Asn Leu Leu Phe
100 105 110

Arg Asn Leu Ile Lys Gly Leu Asn Glu Lys Ile Lys Val Asn Ser Gln
115 120 125

Lys Asp Ile Ile Ile Asp Cys Ile Phe Asp Pro Tyr Asn Glu Gly Arg
130 135 140

Val Val Phe Lys Lys Lys Arg Val Asp Val Ala Ile Ile Leu Lys Asn
145 150 155 160

Lys Phe Val Phe Asn Asn Val Glu Ile Ser Asp Phe Ala Ile Pro Leu
165 170 175

Val Ala Ile Glu Ile Lys Thr Asn Leu Asp Lys Asn Met Leu Ser Gly
180 185 190

Ile Glu Gln Ser Val Asp Ser Leu Lys Glu Thr Phe Pro Leu Cys Leu
195 200 205

```
Tyr Tyr Cys Ile Thr Glu Leu Ala Asp Phe Ala Ile Glu Lys Gln Asn
    210                 215                 220

Tyr Ala Ser Thr His Ile Asp Glu Val Phe Ile Leu Arg Lys Gln Lys
225                 230                 235                 240

Arg Gly Pro Val Arg Arg Gly Thr Pro Leu Glu Val Val His Ala Asp
                245                 250                 255

Leu Ile Leu Glu Val Val Glu Gln Val Gly Glu His Leu Ser Lys Phe
            260                 265                 270

Lys Asp Pro Ile Lys Thr Leu Lys Ala Arg Met Thr Glu Gly Tyr Leu
        275                 280                 285

Ile Lys Gly Lys Gly Lys
    290
```

<210> 53

<211> 288
<212> PRT
<213> Bacillus pumilus

<220>
<223> Bpu10I beta subunit CAA74999.1

<400> 53

```
Met Thr Gln Ile Asp Leu Ser Asn Thr Lys His Gly Ser Ile Leu Phe
1               5                   10                  15

Glu Lys Gln Lys Asn Val Lys Glu Lys Tyr Leu Gln Gln Ala Tyr Lys
            20                  25                  30

His Tyr Leu Tyr Phe Arg Arg Ser Ile Asp Gly Leu Glu Ile Thr Asn
        35                  40                  45

Asp Glu Ala Ile Phe Lys Leu Thr Gln Ala Ala Asn Asn Tyr Arg Asp
    50                  55                  60

Asn Val Leu Tyr Leu Phe Glu Ser Arg Pro Asn Ser Gly Gln Glu Ala
65                  70                  75                  80

Phe Arg Tyr Thr Ile Leu Glu Glu Phe Phe Tyr His Leu Phe Lys Asp
                85                  90                  95

Leu Val Lys Lys Lys Phe Asn Gln Glu Pro Ser Ser Ile Val Met Gly
            100                 105                 110

Lys Ala Asn Ser Tyr Val Ser Leu Ser Phe Ser Pro Glu Ser Phe Leu
        115                 120                 125
```

Gly Leu Tyr Glu Asn Pro Ile Pro Tyr Ile His Thr Lys Asp Gln Asp
130 135 140

Phe Val Leu Gly Cys Ala Val Asp Leu Lys Ile Ser Pro Lys Asn Glu
145 150 155 160

Leu Asn Lys Glu Asn Glu Thr Glu Ile Val Val Pro Val Ile Ala Ile
165 170 175

Glu Cys Lys Thr Tyr Ile Glu Arg Asn Met Leu Asp Ser Cys Ala Ala
180 185 190

Thr Ala Ser Arg Leu Lys Ala Ala Met Pro Tyr Cys Leu Tyr Ile Val
195 200 205

Ala Ser Glu Tyr Met Lys Met Asp Gln Ala Tyr Pro Glu Leu Thr Asp
210 215 220

Ile Asp Glu Val Phe Ile Leu Cys Lys Ala Ser Val Gly Glu Arg Thr
225 230 235 240

Ala Leu Lys Lys Lys Gly Leu Pro Pro His Lys Leu Asp Glu Asn Leu
245 250 255

Met Val Glu Leu Phe His Met Val Glu Arg His Leu Asn Arg Val Trp
260 265 270

Trp Ser Pro Asn Glu Ala Leu Ser Arg Gly Arg Val Ile Gly Arg Pro
275 280 285

<210> 54

<211> 358
<212> PRT
<213> Bacillus megaterium

<220>
<223> BmrI ABM69266.1

<400> 54
Met Asn Tyr Phe Ser Leu His Pro Asn Val Tyr Ala Thr Gly Arg Pro
1 5 10 15

Lys Gly Leu Ile Asn Met Leu Glu Ser Val Trp Ile Ser Asn Gln Lys
20 25 30

Pro Gly Asp Gly Thr Met Tyr Leu Ile Ser Gly Phe Ala Asn Tyr Asn
35 40 45

Gly Gly Ile Arg Phe Tyr Glu Thr Phe Thr Glu His Ile Asn His Gly
50 55 60

Gly Lys Val Ile Ala Ile Leu Gly Gly Ser Thr Ser Gln Arg Leu Ser
65 70 75 80

```
Ser Lys Gln Val Val Ala Glu Leu Val Ser Arg Gly Val Asp Val Tyr
                85                  90                  95

Ile Ile Asn Arg Lys Arg Leu Leu His Ala Lys Leu Tyr Gly Ser Ser
            100                 105                 110

Ser Asn Ser Gly Glu Ser Leu Val Val Ser Ser Gly Asn Phe Thr Gly
        115                 120                 125

Pro Gly Met Ser Gln Asn Val Glu Ala Ser Leu Leu Leu Asp Asn Asn
    130                 135                 140

Thr Thr Ser Ser Met Gly Phe Ser Trp Asn Gly Met Val Asn Ser Met
145                 150                 155                 160

Leu Asp Gln Lys Trp Gln Ile His Asn Leu Ser Asn Ser Asn Pro Thr
                165                 170                 175

Ser Pro Ser Trp Asn Leu Leu Tyr Asp Glu Arg Thr Thr Asn Leu Thr
            180                 185                 190

Leu Asp Asp Thr Gln Lys Val Thr Leu Ile Leu Thr Leu Gly His Ala
        195                 200                 205

Asp Thr Ala Arg Ile Gln Ala Ala Pro Lys Ser Lys Ala Gly Glu Gly
    210                 215                 220

Ser Gln Tyr Phe Trp Leu Ser Lys Asp Ser Tyr Asp Phe Phe Pro Pro
225                 230                 235                 240

Leu Thr Ile Arg Asn Lys Arg Gly Thr Lys Ala Thr Tyr Ser Cys Leu
                245                 250                 255

Ile Asn Met Asn Tyr Leu Asp Ile Lys Tyr Ile Asp Ser Glu Cys Arg
            260                 265                 270

Val Thr Phe Glu Ala Glu Asn Asn Phe Asp Phe Arg Leu Gly Thr Gly
        275                 280                 285

Lys Leu Arg Tyr Thr Asn Val Ala Ala Ser Asp Asp Ile Ala Ala Ile
    290                 295                 300

Thr Arg Val Gly Asp Ser Asp Tyr Glu Leu Arg Ile Ile Lys Lys Gly
305                 310                 315                 320

Ser Ser Asn Tyr Asp Ala Leu Asp Ser Ala Ala Val Asn Phe Ile Gly
                325                 330                 335

Asn Arg Gly Lys Arg Tyr Gly Tyr Ile Pro Asn Asp Glu Phe Gly Arg
            340                 345                 350
```

Ile Ile Gly Ala Lys Phe
355

<210> 55

<211> 358
<212> PRT
<213> Bacillus firmus

<220>
<223> BfiI CAC12783.1

<400> 55
Met Asn Phe Phe Ser Leu His Pro Asn Val Tyr Ala Thr Gly Arg Pro
1 5 10 15

Lys Gly Leu Ile Gly Met Leu Glu Asn Val Trp Val Ser Asn His Thr
20 25 30

Pro Gly Glu Gly Thr Leu Tyr Leu Ile Ser Gly Phe Ser Asn Tyr Asn
35 40 45

Gly Gly Val Arg Phe Tyr Glu Thr Phe Thr Glu His Ile Asn Gln Gly
50 55 60

Gly Arg Val Ile Ala Ile Leu Gly Gly Ser Thr Ser Gln Arg Leu Ser
65 70 75 80

Ser Arg Gln Val Val Glu Glu Leu Leu Asn Arg Gly Val Glu Val His
85 90 95

Ile Ile Asn Arg Lys Arg Ile Leu His Ala Lys Leu Tyr Gly Thr Ser
100 105 110

Asn Asn Leu Gly Glu Ser Leu Val Val Ser Ser Gly Asn Phe Thr Gly
115 120 125

Pro Gly Met Ser Gln Asn Ile Glu Ala Ser Leu Leu Leu Asp Asn Asn
130 135 140

Thr Thr Gln Ser Met Gly Phe Ser Trp Asn Asp Met Ile Ser Glu Met
145 150 155 160

Leu Asn Gln Asn Trp His Ile His Asn Met Thr Asn Ala Thr Asp Ala
165 170 175

Ser Pro Gly Trp Asn Leu Leu Tyr Asp Glu Arg Thr Thr Asn Leu Thr
180 185 190

Leu Asp Glu Thr Glu Arg Val Thr Leu Ile Val Thr Leu Gly His Ala
195 200 205

```
Asp Thr Ala Arg Ile Gln Ala Ala Pro Gly Thr Thr Ala Gly Gln Gly
    210                 215                 220

Thr Gln Tyr Phe Trp Leu Ser Lys Asp Ser Tyr Asp Phe Phe Pro Pro
225                 230                 235                 240

Leu Thr Ile Arg Asn Arg Arg Gly Thr Lys Ala Thr Tyr Ser Ser Leu
                245                 250                 255

Ile Asn Met Asn Tyr Ile Asp Ile Asn Tyr Thr Asp Thr Gln Cys Arg
            260                 265                 270

Val Thr Phe Glu Ala Glu Asn Asn Phe Asp Phe Arg Leu Gly Thr Gly
        275                 280                 285

Lys Leu Arg Tyr Thr Gly Val Ala Lys Ser Asn Asp Ile Ala Ala Ile
    290                 295                 300

Thr Arg Val Gly Asp Ser Asp Tyr Glu Leu Arg Ile Ile Lys Gln Gly
305                 310                 315                 320

Thr Pro Glu His Ser Gln Leu Asp Pro Tyr Ala Val Ser Phe Ile Gly
                325                 330                 335

Asn Arg Gly Lys Arg Phe Gly Tyr Ile Ser Asn Glu Glu Phe Gly Arg
            340                 345                 350

Ile Ile Gly Val Thr Phe
        355
```

<210> 56

<211> 846
<212> PRT
<213> Homo sapiens

<220>
<223> hExoI (EXO1_HUMAN) Q9UQ84.2

<400> 56

```
Met Gly Ile Gln Gly Leu Leu Gln Phe Ile Lys Glu Ala Ser Glu Pro
1               5                   10                  15

Ile His Val Arg Lys Tyr Lys Gly Gln Val Val Ala Val Asp Thr Tyr
            20                  25                  30

Cys Trp Leu His Lys Gly Ala Ile Ala Cys Ala Glu Lys Leu Ala Lys
        35                  40                  45

Gly Glu Pro Thr Asp Arg Tyr Val Gly Phe Cys Met Lys Phe Val Asn
    50                  55                  60

Met Leu Leu Ser His Gly Ile Lys Pro Ile Leu Val Phe Asp Gly Cys
65                  70                  75                  80
```

```
Thr Leu Pro Ser Lys Lys Glu Val Glu Arg Ser Arg Arg Glu Arg Arg
                85                  90                  95

Gln Ala Asn Leu Leu Lys Gly Lys Gln Leu Leu Arg Glu Gly Lys Val
            100                 105                 110

Ser Glu Ala Arg Glu Cys Phe Thr Arg Ser Ile Asn Ile Thr His Ala
        115                 120                 125

Met Ala His Lys Val Ile Lys Ala Ala Arg Ser Gln Gly Val Asp Cys
    130                 135                 140

Leu Val Ala Pro Tyr Glu Ala Asp Ala Gln Leu Ala Tyr Leu Asn Lys
145                 150                 155                 160

Ala Gly Ile Val Gln Ala Ile Ile Thr Glu Asp Ser Asp Leu Leu Ala
                165                 170                 175

Phe Gly Cys Lys Lys Val Ile Leu Lys Met Asp Gln Phe Gly Asn Gly
            180                 185                 190

Leu Glu Ile Asp Gln Ala Arg Leu Gly Met Cys Arg Gln Leu Gly Asp
        195                 200                 205

Val Phe Thr Glu Glu Lys Phe Arg Tyr Met Cys Ile Leu Ser Gly Cys
    210                 215                 220

Asp Tyr Leu Ser Ser Leu Arg Gly Ile Gly Leu Ala Lys Ala Cys Lys
225                 230                 235                 240

Val Leu Arg Leu Ala Asn Asn Pro Asp Ile Val Lys Val Ile Lys Lys
                245                 250                 255

Ile Gly His Tyr Leu Lys Met Asn Ile Thr Val Pro Glu Asp Tyr Ile
            260                 265                 270

Asn Gly Phe Ile Arg Ala Asn Asn Thr Phe Leu Tyr Gln Leu Val Phe
        275                 280                 285

Asp Pro Ile Lys Arg Lys Leu Ile Pro Leu Asn Ala Tyr Glu Asp Asp
    290                 295                 300

Val Asp Pro Glu Thr Leu Ser Tyr Ala Gly Gln Tyr Val Asp Asp Ser
305                 310                 315                 320

Ile Ala Leu Gln Ile Ala Leu Gly Asn Lys Asp Ile Asn Thr Phe Glu
                325                 330                 335

Gln Ile Asp Asp Tyr Asn Pro Asp Thr Ala Met Pro Ala His Ser Arg
            340                 345                 350
```

```
Ser His Ser Trp Asp Asp Lys Thr Cys Gln Lys Ser Ala Asn Val Ser
        355                 360                 365

Ser Ile Trp His Arg Asn Tyr Ser Pro Arg Pro Glu Ser Gly Thr Val
    370                 375                 380

Ser Asp Ala Pro Gln Leu Lys Glu Asn Pro Ser Thr Val Gly Val Glu
385                 390                 395                 400

Arg Val Ile Ser Thr Lys Gly Leu Asn Leu Pro Arg Lys Ser Ser Ile
                405                 410                 415

Val Lys Arg Pro Arg Ser Ala Glu Leu Ser Glu Asp Asp Leu Leu Ser
            420                 425                 430

Gln Tyr Ser Leu Ser Phe Thr Lys Lys Thr Lys Lys Asn Ser Ser Glu
        435                 440                 445

Gly Asn Lys Ser Leu Ser Phe Ser Glu Val Phe Val Pro Asp Leu Val
    450                 455                 460

Asn Gly Pro Thr Asn Lys Lys Ser Val Ser Thr Pro Pro Arg Thr Arg
465                 470                 475                 480

Asn Lys Phe Ala Thr Phe Leu Gln Arg Lys Asn Glu Glu Ser Gly Ala
                485                 490                 495

Val Val Val Pro Gly Thr Arg Ser Arg Phe Phe Cys Ser Ser Asp Ser
            500                 505                 510

Thr Asp Cys Val Ser Asn Lys Val Ser Ile Gln Pro Leu Asp Glu Thr
        515                 520                 525

Ala Val Thr Asp Lys Glu Asn Asn Leu His Glu Ser Glu Tyr Gly Asp
    530                 535                 540

Gln Glu Gly Lys Arg Leu Val Asp Thr Asp Val Ala Arg Asn Ser Ser
545                 550                 555                 560

Asp Asp Ile Pro Asn Asn His Ile Pro Gly Asp His Ile Pro Asp Lys
                565                 570                 575

Ala Thr Val Phe Thr Asp Glu Glu Ser Tyr Ser Phe Glu Ser Ser Lys
            580                 585                 590

Phe Thr Arg Thr Ile Ser Pro Pro Thr Leu Gly Thr Leu Arg Ser Cys
        595                 600                 605

Phe Ser Trp Ser Gly Gly Leu Gly Asp Phe Ser Arg Thr Pro Ser Pro
    610                 615                 620
```

```
Ser Pro Ser Thr Ala Leu Gln Gln Phe Arg Arg Lys Ser Asp Ser Pro
625                 630                 635                 640

Thr Ser Leu Pro Glu Asn Asn Met Ser Asp Val Ser Gln Leu Lys Ser
                645                 650                 655

Glu Glu Ser Ser Asp Asp Glu Ser His Pro Leu Arg Glu Glu Ala Cys
            660                 665                 670

Ser Ser Gln Ser Gln Glu Ser Gly Glu Phe Ser Leu Gln Ser Ser Asn
        675                 680                 685

Ala Ser Lys Leu Ser Gln Cys Ser Ser Lys Asp Ser Asp Ser Glu Glu
    690                 695                 700

Ser Asp Cys Asn Ile Lys Leu Leu Asp Ser Gln Ser Asp Gln Thr Ser
705                 710                 715                 720

Lys Leu Arg Leu Ser His Phe Ser Lys Lys Asp Thr Pro Leu Arg Asn
                725                 730                 735

Lys Val Pro Gly Leu Tyr Lys Ser Ser Ser Ala Asp Ser Leu Ser Thr
            740                 745                 750

Thr Lys Ile Lys Pro Leu Gly Pro Ala Arg Ala Ser Gly Leu Ser Lys
        755                 760                 765

Lys Pro Ala Ser Ile Gln Lys Arg Lys His His Asn Ala Glu Asn Lys
    770                 775                 780

Pro Gly Leu Gln Ile Lys Leu Asn Glu Leu Trp Lys Asn Phe Gly Phe
785                 790                 795                 800

Lys Lys Asp Ser Glu Lys Leu Pro Pro Cys Lys Lys Pro Leu Ser Pro
                805                 810                 815

Val Arg Asp Asn Ile Gln Leu Thr Pro Glu Ala Glu Glu Asp Ile Phe
            820                 825                 830

Asn Lys Pro Glu Cys Gly Arg Val Gln Arg Ala Ile Phe Gln
        835                 840                 845
```

<210> 57

<211> 702
<212> PRT
<213> Saccharomyces cerevisiae

<220>
<223> Yeast ExoI (EXO1_YEAST) P39875.2

<400> 57

```
Met Gly Ile Gln Gly Leu Leu Pro Gln Leu Lys Pro Ile Gln Asn Pro
1               5                  10                  15

Val Ser Leu Arg Arg Tyr Glu Gly Glu Val Leu Ala Ile Asp Gly Tyr
            20                  25                  30

Ala Trp Leu His Arg Ala Ala Cys Ser Cys Ala Tyr Glu Leu Ala Met
        35                  40                  45

Gly Lys Pro Thr Asp Lys Tyr Leu Gln Phe Phe Ile Lys Arg Phe Ser
    50                  55                  60

Leu Leu Lys Thr Phe Lys Val Glu Pro Tyr Leu Val Phe Asp Gly Asp
65                  70                  75                  80

Ala Ile Pro Val Lys Lys Ser Thr Glu Ser Lys Arg Arg Asp Lys Arg
                85                  90                  95

Lys Glu Asn Lys Ala Ile Ala Glu Arg Leu Trp Ala Cys Gly Glu Lys
            100                 105                 110

Lys Asn Ala Met Asp Tyr Phe Gln Lys Cys Val Asp Ile Thr Pro Glu
        115                 120                 125

Met Ala Lys Cys Ile Ile Cys Tyr Cys Lys Leu Asn Gly Ile Arg Tyr
    130                 135                 140

Ile Val Ala Pro Phe Glu Ala Asp Ser Gln Met Val Tyr Leu Glu Gln
145                 150                 155                 160

Lys Asn Ile Val Gln Gly Ile Ile Ser Glu Asp Ser Asp Leu Leu Val
                165                 170                 175

Phe Gly Cys Arg Arg Leu Ile Thr Lys Leu Asn Asp Tyr Gly Glu Cys
            180                 185                 190

Leu Glu Ile Cys Arg Asp Asn Phe Ile Lys Leu Pro Lys Lys Phe Pro
        195                 200                 205

Leu Gly Ser Leu Thr Asn Glu Glu Ile Ile Thr Met Val Cys Leu Ser
    210                 215                 220

Gly Cys Asp Tyr Thr Asn Gly Ile Pro Lys Val Gly Leu Ile Thr Ala
225                 230                 235                 240

Met Lys Leu Val Arg Arg Phe Asn Thr Ile Glu Arg Ile Ile Leu Ser
                245                 250                 255

Ile Gln Arg Glu Gly Lys Leu Met Ile Pro Asp Thr Tyr Ile Asn Glu
            260                 265                 270
```

```
Tyr Glu Ala Ala Val Leu Ala Phe Gln Phe Gln Arg Val Phe Cys Pro
        275                 280                 285

Ile Arg Lys Lys Ile Val Ser Leu Asn Glu Ile Pro Leu Tyr Leu Lys
    290                 295                 300

Asp Thr Glu Ser Lys Arg Lys Arg Leu Tyr Ala Cys Ile Gly Phe Val
305                 310                 315                 320

Ile His Arg Glu Thr Gln Lys Lys Gln Ile Val His Phe Asp Asp Asp
                325                 330                 335

Ile Asp His His Leu His Leu Lys Ile Ala Gln Gly Asp Leu Asn Pro
            340                 345                 350

Tyr Asp Phe His Gln Pro Leu Ala Asn Arg Glu His Lys Leu Gln Leu
        355                 360                 365

Ala Ser Lys Ser Asn Ile Glu Phe Gly Lys Thr Asn Thr Thr Asn Ser
    370                 375                 380

Glu Ala Lys Val Lys Pro Ile Glu Ser Phe Phe Gln Lys Met Thr Lys
385                 390                 395                 400

Leu Asp His Asn Pro Lys Val Ala Asn Asn Ile His Ser Leu Arg Gln
                405                 410                 415

Ala Glu Asp Lys Leu Thr Met Ala Ile Lys Arg Arg Lys Leu Ser Asn
            420                 425                 430

Ala Asn Val Val Gln Glu Thr Leu Lys Asp Thr Arg Ser Lys Phe Phe
        435                 440                 445

Asn Lys Pro Ser Met Thr Val Val Glu Asn Phe Lys Glu Lys Gly Asp
    450                 455                 460

Ser Ile Gln Asp Phe Lys Glu Asp Thr Asn Ser Gln Ser Leu Glu Glu
465                 470                 475                 480

Pro Val Ser Glu Ser Gln Leu Ser Thr Gln Ile Pro Ser Ser Phe Ile
                485                 490                 495

Thr Thr Asn Leu Glu Asp Asp Asp Asn Leu Ser Glu Glu Val Ser Glu
            500                 505                 510

Val Val Ser Asp Ile Glu Glu Asp Arg Lys Asn Ser Glu Gly Lys Thr
        515                 520                 525

Ile Gly Asn Glu Ile Tyr Asn Thr Asp Asp Asp Gly Asp Gly Asp Thr
    530                 535                 540
```

Ser Glu Asp Tyr Ser Glu Thr Ala Glu Ser Arg Val Pro Thr Ser Ser
545 550 555 560

Thr Thr Ser Phe Pro Gly Ser Ser Gln Arg Ser Ile Ser Gly Cys Thr
565 570 575

Lys Val Leu Gln Lys Phe Arg Tyr Ser Ser Ser Phe Ser Gly Val Asn
580 585 590

Ala Asn Arg Gln Pro Leu Phe Pro Arg His Val Asn Gln Lys Ser Arg
595 600 605

Gly Met Val Tyr Val Asn Gln Asn Arg Asp Asp Asp Cys Asp Asp Asn
610 615 620

Asp Gly Lys Asn Gln Ile Thr Gln Arg Pro Ser Leu Arg Lys Ser Leu
625 630 635 640

Ile Gly Ala Arg Ser Gln Arg Ile Val Ile Asp Met Lys Ser Val Asp
645 650 655

Glu Arg Lys Ser Phe Asn Ser Ser Pro Ile Leu His Glu Glu Ser Lys
660 665 670

Lys Arg Asp Ile Glu Thr Thr Lys Ser Ser Gln Ala Arg Pro Ala Val
675 680 685

Arg Ser Ile Ser Leu Leu Ser Gln Phe Val Tyr Lys Gly Lys
690 695 700

<210> 58

<211> 475
<212> PRT
<213> Escherichia coli DH1

<220>
<223> E.coli ExoI BAJ43803.1

<400> 58
Met Met Asn Asp Gly Lys Gln Gln Ser Thr Phe Leu Phe His Asp Tyr
1 5 10 15

Glu Thr Phe Gly Thr His Pro Ala Leu Asp Arg Pro Ala Gln Phe Ala
20 25 30

Ala Ile Arg Thr Asp Ser Glu Phe Asn Val Ile Gly Glu Pro Glu Val
35 40 45

Phe Tyr Cys Lys Pro Ala Asp Asp Tyr Leu Pro Gln Pro Gly Ala Val
50 55 60

Leu Ile Thr Gly Ile Thr Pro Gln Glu Ala Arg Ala Lys Gly Glu Asn
65 70 75 80

```
Glu Ala Ala Phe Ala Ala Arg Ile His Ser Leu Phe Thr Val Pro Lys
                85                  90                  95

Thr Cys Ile Leu Gly Tyr Asn Asn Val Arg Phe Asp Asp Glu Val Thr
            100                 105                 110

Arg Asn Ile Phe Tyr Arg Asn Phe Tyr Asp Pro Tyr Ala Trp Ser Trp
        115                 120                 125

Gln His Asp Asn Ser Arg Trp Asp Leu Leu Asp Val Met Arg Ala Cys
    130                 135                 140

Tyr Ala Leu Arg Pro Glu Gly Ile Asn Trp Pro Glu Asn Asp Asp Gly
145                 150                 155                 160

Leu Pro Ser Phe Arg Leu Glu His Leu Thr Lys Ala Asn Gly Ile Glu
                165                 170                 175

His Ser Asn Ala His Asp Ala Met Ala Asp Val Tyr Ala Thr Ile Ala
            180                 185                 190

Met Ala Lys Leu Val Lys Thr Arg Gln Pro Arg Leu Phe Asp Tyr Leu
        195                 200                 205

Phe Thr His Arg Asn Lys His Lys Leu Met Ala Leu Ile Asp Val Pro
    210                 215                 220

Gln Met Lys Pro Leu Val His Val Ser Gly Met Phe Gly Ala Trp Arg
225                 230                 235                 240

Gly Asn Thr Ser Trp Val Ala Pro Leu Ala Trp His Pro Glu Asn Arg
                245                 250                 255

Asn Ala Val Ile Met Val Asp Leu Ala Gly Asp Ile Ser Pro Leu Leu
            260                 265                 270

Glu Leu Asp Ser Asp Thr Leu Arg Glu Arg Leu Tyr Thr Ala Lys Thr
        275                 280                 285

Asp Leu Gly Asp Asn Ala Ala Val Pro Val Lys Leu Val His Ile Asn
    290                 295                 300

Lys Cys Pro Val Leu Ala Gln Ala Asn Thr Leu Arg Pro Glu Asp Ala
305                 310                 315                 320

Asp Arg Leu Gly Ile Asn Arg Gln His Cys Leu Asp Asn Leu Lys Ile
                325                 330                 335

Leu Arg Glu Asn Pro Gln Val Arg Glu Lys Val Val Ala Ile Phe Ala
            340                 345                 350
```

Glu Ala Glu Pro Phe Thr Pro Ser Asp Asn Val Asp Ala Gln Leu Tyr
355 360 365

Asn Gly Phe Phe Ser Asp Ala Asp Arg Ala Ala Met Lys Ile Val Leu
370 375 380

Glu Thr Glu Pro Arg Asn Leu Pro Ala Leu Asp Ile Thr Phe Val Asp
385 390 395 400

Lys Arg Ile Glu Lys Leu Leu Phe Asn Tyr Arg Ala Arg Asn Phe Pro
405 410 415

Gly Thr Leu Asp Tyr Ala Glu Gln Gln Arg Trp Leu Glu His Arg Arg
420 425 430

Gln Val Phe Thr Pro Glu Phe Leu Gln Gly Tyr Ala Asp Glu Leu Gln
435 440 445

Met Leu Val Gln Gln Tyr Ala Asp Asp Lys Glu Lys Val Ala Leu Leu
450 455 460

Lys Ala Leu Trp Gln Tyr Ala Glu Glu Ile Val
465 470 475

<210> 59

<211> 279
<212> PRT
<213> Homo sapiens

<220>
<223> TREX2_HUMAN Q9BQ50.1

<400> 59
Met Gly Arg Ala Gly Ser Pro Leu Pro Arg Ser Ser Trp Pro Arg Met
1 5 10 15

Asp Asp Cys Gly Ser Arg Ser Arg Cys Ser Pro Thr Leu Cys Ser Ser
20 25 30

Leu Arg Thr Cys Tyr Pro Arg Gly Asn Ile Thr Met Ser Glu Ala Pro
35 40 45

Arg Ala Glu Thr Phe Val Phe Leu Asp Leu Glu Ala Thr Gly Leu Pro
50 55 60

Ser Val Glu Pro Glu Ile Ala Glu Leu Ser Leu Phe Ala Val His Arg
65 70 75 80

Ser Ser Leu Glu Asn Pro Glu His Asp Glu Ser Gly Ala Leu Val Leu
85 90 95

Pro Arg Val Leu Asp Lys Leu Thr Leu Cys Met Cys Pro Glu Arg Pro
100 105 110

Phe Thr Ala Lys Ala Ser Glu Ile Thr Gly Leu Ser Ser Glu Gly Leu
115 120 125

Ala Arg Cys Arg Lys Ala Gly Phe Asp Gly Ala Val Val Arg Thr Leu
130 135 140

Gln Ala Phe Leu Ser Arg Gln Ala Gly Pro Ile Cys Leu Val Ala His
145 150 155 160

Asn Gly Phe Asp Tyr Asp Phe Pro Leu Leu Cys Ala Glu Leu Arg Arg
165 170 175

Leu Gly Ala Arg Leu Pro Arg Asp Thr Val Cys Leu Asp Thr Leu Pro
180 185 190

Ala Leu Arg Gly Leu Asp Arg Ala His Ser His Gly Thr Arg Ala Arg
195 200 205

Gly Arg Gln Gly Tyr Ser Leu Gly Ser Leu Phe His Arg Tyr Phe Arg
210 215 220

Ala Glu Pro Ser Ala Ala His Ser Ala Glu Gly Asp Val His Thr Leu
225 230 235 240

Leu Leu Ile Phe Leu His Arg Ala Ala Glu Leu Leu Ala Trp Ala Asp
245 250 255

Glu Gln Ala Arg Gly Trp Ala His Ile Glu Pro Met Tyr Leu Pro Pro
260 265 270

Asp Asp Pro Ser Leu Glu Ala
275

<210> 60

<211> 314
<212> PRT
<213> Mus musculus

<220>
<223> TREX1_MOUSE Q91XB0.2

<400> 60
Met Gly Ser Gln Thr Leu Pro His Gly His Met Gln Thr Leu Ile Phe
1 5 10 15

Leu Asp Leu Glu Ala Thr Gly Leu Pro Ser Ser Arg Pro Glu Val Thr
20 25 30

Glu Leu Cys Leu Leu Ala Val His Arg Arg Ala Leu Glu Asn Thr Ser
35 40 45

```
Ile Ser Gln Gly His Pro Pro Pro Val Pro Arg Pro Pro Arg Val Val
    50                  55                  60

Asp Lys Leu Ser Leu Cys Ile Ala Pro Gly Lys Ala Cys Ser Pro Gly
65                  70                  75                  80

Ala Ser Glu Ile Thr Gly Leu Ser Lys Ala Glu Leu Glu Val Gln Gly
                85                  90                  95

Arg Gln Arg Phe Asp Asp Asn Leu Ala Ile Leu Leu Arg Ala Phe Leu
            100                 105                 110

Gln Arg Gln Pro Gln Pro Cys Cys Leu Val Ala His Asn Gly Asp Arg
        115                 120                 125

Tyr Asp Phe Pro Leu Leu Gln Thr Glu Leu Ala Arg Leu Ser Thr Pro
    130                 135                 140

Ser Pro Leu Asp Gly Thr Phe Cys Val Asp Ser Ile Ala Ala Leu Lys
145                 150                 155                 160

Ala Leu Glu Gln Ala Ser Ser Pro Ser Gly Asn Gly Ser Arg Lys Ser
                165                 170                 175

Tyr Ser Leu Gly Ser Ile Tyr Thr Arg Leu Tyr Trp Gln Ala Pro Thr
            180                 185                 190

Asp Ser His Thr Ala Glu Gly Asp Val Leu Thr Leu Leu Ser Ile Cys
        195                 200                 205

Gln Trp Lys Pro Gln Ala Leu Leu Gln Trp Val Asp Glu His Ala Arg
    210                 215                 220

Pro Phe Ser Thr Val Lys Pro Met Tyr Gly Thr Pro Ala Thr Thr Gly
225                 230                 235                 240

Thr Thr Asn Leu Arg Pro His Ala Ala Thr Ala Thr Thr Pro Leu Ala
                245                 250                 255

Thr Ala Asn Gly Ser Pro Ser Asn Gly Arg Ser Arg Arg Pro Lys Ser
            260                 265                 270

Pro Pro Pro Glu Lys Val Pro Glu Ala Pro Ser Gln Glu Gly Leu Leu
        275                 280                 285

Ala Pro Leu Ser Leu Leu Thr Leu Leu Thr Leu Ala Ile Ala Thr Leu
    290                 295                 300

Tyr Gly Leu Phe Leu Ala Ser Pro Gly Gln
305                 310
```

<210> 61

<211> 369
<212> PRT
<213> Homo sapiens

<220>
<223> TREX1_HUMAN Q9NSU2.1

<400> 61
Met Gly Pro Gly Ala Arg Arg Gln Gly Arg Ile Val Gln Gly Arg Pro
1 5 10 15

Glu Met Cys Phe Cys Pro Pro Pro Thr Pro Leu Pro Pro Leu Arg Ile
20 25 30

Leu Thr Leu Gly Thr His Thr Pro Thr Pro Cys Ser Ser Pro Gly Ser
35 40 45

Ala Ala Gly Thr Tyr Pro Thr Met Gly Ser Gln Ala Leu Pro Pro Gly
50 55 60

Pro Met Gln Thr Leu Ile Phe Phe Asp Met Glu Ala Thr Gly Leu Pro
65 70 75 80

Phe Ser Gln Pro Lys Val Thr Glu Leu Cys Leu Leu Ala Val His Arg
85 90 95

Cys Ala Leu Glu Ser Pro Pro Thr Ser Gln Gly Pro Pro Pro Thr Val
100 105 110

Pro Pro Pro Pro Arg Val Val Asp Lys Leu Ser Leu Cys Val Ala Pro
115 120 125

Gly Lys Ala Cys Ser Pro Ala Ala Ser Glu Ile Thr Gly Leu Ser Thr
130 135 140

Ala Val Leu Ala Ala His Gly Arg Gln Cys Phe Asp Asp Asn Leu Ala
145 150 155 160

Asn Leu Leu Leu Ala Phe Leu Arg Arg Gln Pro Gln Pro Trp Cys Leu
165 170 175

Val Ala His Asn Gly Asp Arg Tyr Asp Phe Pro Leu Leu Gln Ala Glu
180 185 190

Leu Ala Met Leu Gly Leu Thr Ser Ala Leu Asp Gly Ala Phe Cys Val
195 200 205

Asp Ser Ile Thr Ala Leu Lys Ala Leu Glu Arg Ala Ser Ser Pro Ser
210 215 220

```
Glu His Gly Pro Arg Lys Ser Tyr Ser Leu Gly Ser Ile Tyr Thr Arg
225                 230                 235                 240

Leu Tyr Gly Gln Ser Pro Pro Asp Ser His Thr Ala Glu Gly Asp Val
                245                 250                 255

Leu Ala Leu Leu Ser Ile Cys Gln Trp Arg Pro Gln Ala Leu Leu Arg
            260                 265                 270

Trp Val Asp Ala His Ala Arg Pro Phe Gly Thr Ile Arg Pro Met Tyr
        275                 280                 285

Gly Val Thr Ala Ser Ala Arg Thr Lys Pro Arg Pro Ser Ala Val Thr
    290                 295                 300

Thr Thr Ala His Leu Ala Thr Thr Arg Asn Thr Ser Pro Ser Leu Gly
305                 310                 315                 320

Glu Ser Arg Gly Thr Lys Asp Leu Pro Pro Val Lys Asp Pro Gly Ala
                325                 330                 335

Leu Ser Arg Glu Gly Leu Leu Ala Pro Leu Gly Leu Leu Ala Ile Leu
            340                 345                 350

Thr Leu Ala Val Ala Thr Leu Tyr Gly Leu Ser Leu Ala Thr Pro Gly
        355                 360                 365

Glu
```

<210> 62

<211> 315
<212> PRT
<213> Bos taurus

<220>
<223> TREX1_BOVIN Q9BG99.1

<400> 62

```
Met Gly Ser Arg Ala Leu Pro Pro Gly Pro Val Gln Thr Leu Ile Phe
1               5                   10                  15

Leu Asp Leu Glu Ala Thr Gly Leu Pro Phe Ser Gln Pro Lys Ile Thr
            20                  25                  30

Glu Leu Cys Leu Leu Ala Val His Arg Tyr Ala Leu Glu Gly Leu Ser
        35                  40                  45

Ala Pro Gln Gly Pro Ser Pro Thr Ala Pro Val Pro Pro Arg Val Leu
    50                  55                  60

Asp Lys Leu Ser Leu Cys Val Ala Pro Gly Lys Val Cys Ser Pro Ala
65                  70                  75                  80
```

```
Ala Ser Glu Ile Thr Gly Leu Ser Thr Ala Val Leu Ala Ala His Gly
                85                  90                  95

Arg Arg Ala Phe Asp Ala Asp Leu Val Asn Leu Ile Arg Thr Phe Leu
            100                 105                 110

Gln Arg Gln Pro Gln Pro Trp Cys Leu Val Ala His Asn Gly Asp Arg
        115                 120                 125

Tyr Asp Phe Pro Leu Leu Arg Ala Glu Leu Ala Leu Leu Gly Leu Ala
    130                 135                 140

Ser Ala Leu Asp Asp Ala Phe Cys Val Asp Ser Ile Ala Ala Leu Lys
145                 150                 155                 160

Ala Leu Glu Pro Thr Gly Ser Ser Ser Glu His Gly Pro Arg Lys Ser
                165                 170                 175

Tyr Ser Leu Gly Ser Val Tyr Thr Arg Leu Tyr Gly Gln Ala Pro Pro
            180                 185                 190

Asp Ser His Thr Ala Glu Gly Asp Val Leu Ala Leu Leu Ser Val Cys
        195                 200                 205

Gln Trp Arg Pro Arg Ala Leu Leu Arg Trp Val Asp Ala His Ala Lys
    210                 215                 220

Pro Phe Ser Thr Val Lys Pro Met Tyr Val Ile Thr Thr Ser Thr Gly
225                 230                 235                 240

Thr Asn Pro Arg Pro Ser Ala Val Thr Ala Thr Val Pro Leu Ala Arg
                245                 250                 255

Ala Ser Asp Thr Gly Pro Asn Leu Arg Gly Asp Arg Ser Pro Lys Pro
            260                 265                 270

Ala Pro Ser Pro Lys Met Cys Pro Gly Ala Pro Pro Gly Glu Gly Leu
        275                 280                 285

Leu Ala Pro Leu Gly Leu Leu Ala Phe Leu Thr Leu Ala Val Ala Met
    290                 295                 300

Leu Tyr Gly Leu Ser Leu Ala Met Pro Gly Gln
305                 310                 315
```

<210> 63

<211> 316
<212> PRT
<213> Rattus norvegicus

<220>

<223> Rat TREX1 AAH91242.1

<400> 63

```
Met Gly Ser Gln Ala Leu Pro His Gly His Met Gln Thr Leu Ile Phe
1               5                  10                  15

Leu Asp Leu Glu Ala Thr Gly Leu Pro Tyr Ser Gln Pro Lys Ile Thr
            20                  25                  30

Glu Leu Cys Leu Leu Ala Val His Arg His Ala Leu Glu Asn Ser Ser
        35                  40                  45

Met Ser Glu Gly Gln Pro Pro Pro Val Pro Lys Pro Pro Arg Val Val
    50                  55                  60

Asp Lys Leu Ser Leu Cys Ile Ala Pro Gly Lys Pro Cys Ser Ser Gly
65                  70                  75                  80

Ala Ser Glu Ile Thr Gly Leu Thr Thr Ala Gly Leu Glu Ala His Gly
                85                  90                  95

Arg Gln Arg Phe Asn Asp Asn Leu Ala Thr Leu Leu Gln Val Phe Leu
            100                 105                 110

Gln Arg Gln Pro Gln Pro Cys Cys Leu Val Ala His Asn Gly Asp Arg
        115                 120                 125

Tyr Asp Phe Pro Leu Leu Gln Ala Glu Leu Ala Ser Leu Ser Val Ile
    130                 135                 140

Ser Pro Leu Asp Gly Thr Phe Cys Val Asp Ser Ile Ala Ala Leu Lys
145                 150                 155                 160

Thr Leu Glu Gln Ala Ser Ser Pro Ser Glu His Gly Pro Arg Lys Ser
                165                 170                 175

Tyr Ser Leu Gly Ser Ile Tyr Thr Arg Leu Tyr Gly Gln Ala Pro Thr
            180                 185                 190

Asp Ser His Thr Ala Glu Gly Asp Val Leu Ala Leu Leu Ser Ile Cys
        195                 200                 205

Gln Trp Lys Pro Gln Ala Leu Leu Gln Trp Val Asp Lys His Ala Arg
    210                 215                 220

Pro Phe Ser Thr Ile Lys Pro Met Tyr Gly Met Ala Ala Thr Thr Gly
225                 230                 235                 240

Thr Ala Ser Pro Arg Leu Cys Ala Ala Thr Thr Ser Ser Pro Leu Ala
                245                 250                 255

Thr Ala Asn Leu Ser Pro Ser Asn Gly Arg Ser Arg Gly Lys Arg Pro
```

260 265 270

Thr Ser Pro Pro Pro Glu Asn Val Pro Glu Ala Pro Ser Arg Glu Gly
275 280 285

Leu Leu Ala Pro Leu Gly Leu Leu Thr Phe Leu Thr Leu Ala Ile Ala
290 295 300

Val Leu Tyr Gly Ile Phe Leu Ala Ser Pro Gly Gln
305 310 315

<210> 64

<211> 829
<212> PRT
<213> Homo sapiens

<220>
<223> Human DNA2 AAH63664.1

<400> 64
Phe Ala Ile Pro Ala Ser Arg Met Glu Gln Leu Asn Glu Leu Glu Leu
1 5 10 15

Leu Met Glu Lys Ser Phe Trp Glu Glu Ala Glu Leu Pro Ala Glu Leu
20 25 30

Phe Gln Lys Lys Val Val Ala Ser Phe Pro Arg Thr Val Leu Ser Thr
35 40 45

Gly Met Asp Asn Arg Tyr Leu Val Leu Ala Val Asn Thr Val Gln Asn
50 55 60

Lys Glu Gly Asn Cys Glu Lys Arg Leu Val Ile Thr Ala Ser Gln Ser
65 70 75 80

Leu Glu Asn Lys Glu Leu Cys Ile Leu Arg Asn Asp Trp Cys Ser Val
85 90 95

Pro Val Glu Pro Gly Asp Ile Ile His Leu Glu Gly Asp Cys Thr Ser
100 105 110

Asp Thr Trp Ile Ile Asp Lys Asp Phe Gly Tyr Leu Ile Leu Tyr Pro
115 120 125

Asp Met Leu Ile Ser Gly Thr Ser Ile Ala Ser Ser Ile Arg Cys Met
130 135 140

Arg Arg Ala Val Leu Ser Glu Thr Phe Arg Ser Ser Asp Pro Ala Thr
145 150 155 160

Arg Gln Met Leu Ile Gly Thr Val Leu His Glu Val Phe Gln Lys Ala
165 170 175

```
Ile Asn Asn Ser Phe Ala Pro Glu Lys Leu Gln Glu Leu Ala Phe Gln
            180                 185                 190

Thr Ile Gln Glu Ile Arg His Leu Lys Glu Met Tyr Arg Leu Asn Leu
        195                 200                 205

Ser Gln Asp Glu Ile Lys Gln Glu Val Glu Asp Tyr Leu Pro Ser Phe
    210                 215                 220

Cys Lys Trp Ala Gly Asp Phe Met His Lys Asn Thr Ser Thr Asp Phe
225                 230                 235                 240

Pro Gln Met Gln Leu Ser Leu Pro Ser Asp Asn Ser Lys Asp Asn Ser
                245                 250                 255

Thr Cys Asn Ile Glu Val Val Lys Pro Met Asp Ile Glu Glu Ser Ile
            260                 265                 270

Trp Ser Pro Arg Phe Gly Leu Lys Gly Lys Ile Asp Val Thr Val Gly
        275                 280                 285

Val Lys Ile His Arg Gly Tyr Lys Thr Lys Tyr Lys Ile Met Pro Leu
    290                 295                 300

Glu Leu Lys Thr Gly Lys Glu Ser Asn Ser Ile Glu His Arg Ser Gln
305                 310                 315                 320

Val Val Leu Tyr Thr Leu Leu Ser Gln Glu Arg Arg Ala Asp Pro Glu
                325                 330                 335

Ala Gly Leu Leu Leu Tyr Leu Lys Thr Gly Gln Met Tyr Pro Val Pro
            340                 345                 350

Ala Asn His Leu Asp Lys Arg Glu Leu Leu Lys Leu Arg Asn Gln Met
        355                 360                 365

Ala Phe Ser Leu Phe His Arg Ile Ser Lys Ser Ala Thr Arg Gln Lys
    370                 375                 380

Thr Gln Leu Ala Ser Leu Pro Gln Ile Ile Glu Glu Glu Lys Thr Cys
385                 390                 395                 400

Lys Tyr Cys Ser Gln Ile Gly Asn Cys Ala Leu Tyr Ser Arg Ala Val
                405                 410                 415

Glu Gln Gln Met Asp Cys Ser Ser Val Pro Ile Val Met Leu Pro Lys
            420                 425                 430

Ile Glu Glu Glu Thr Gln His Leu Lys Gln Thr His Leu Glu Tyr Phe
        435                 440                 445
```

```
Ser Leu Trp Cys Leu Met Leu Thr Leu Glu Ser Gln Ser Lys Asp Asn
    450                 455                 460

Lys Lys Asn His Gln Asn Ile Trp Leu Met Pro Ala Ser Glu Met Glu
465                 470                 475                 480

Lys Ser Gly Ser Cys Ile Gly Asn Leu Ile Arg Met Glu His Val Lys
                485                 490                 495

Ile Val Cys Asp Gly Gln Tyr Leu His Asn Phe Gln Cys Lys His Gly
            500                 505                 510

Ala Ile Pro Val Thr Asn Leu Met Ala Gly Asp Arg Val Ile Val Ser
        515                 520                 525

Gly Glu Glu Arg Ser Leu Phe Ala Leu Ser Arg Gly Tyr Val Lys Glu
    530                 535                 540

Ile Asn Met Thr Thr Val Thr Cys Leu Leu Asp Arg Asn Leu Ser Val
545                 550                 555                 560

Leu Pro Glu Ser Thr Leu Phe Arg Leu Asp Gln Glu Glu Lys Asn Cys
                565                 570                 575

Asp Ile Asp Thr Pro Leu Gly Asn Leu Ser Lys Leu Met Glu Asn Thr
            580                 585                 590

Phe Val Ser Lys Lys Leu Arg Asp Leu Ile Ile Asp Phe Arg Glu Pro
        595                 600                 605

Gln Phe Ile Ser Tyr Leu Ser Ser Val Leu Pro His Asp Ala Lys Asp
    610                 615                 620

Thr Val Ala Cys Ile Leu Lys Gly Leu Asn Lys Pro Gln Arg Gln Ala
625                 630                 635                 640

Met Lys Lys Val Leu Leu Ser Lys Asp Tyr Thr Leu Ile Val Gly Met
                645                 650                 655

Pro Gly Thr Gly Lys Thr Thr Thr Ile Cys Thr Leu Val Pro Ala Pro
            660                 665                 670

Glu Gln Val Glu Lys Gly Gly Val Ser Asn Val Thr Glu Ala Lys Leu
        675                 680                 685

Ile Val Phe Leu Thr Ser Ile Phe Val Lys Ala Gly Cys Ser Pro Ser
    690                 695                 700

Asp Ile Gly Ile Ile Ala Pro Tyr Arg Gln Gln Leu Lys Ile Ile Asn
705                 710                 715                 720
```

```
Asp Leu Leu Ala Arg Ser Ile Gly Met Val Glu Val Asn Thr Val Asp
                725                 730                 735

Lys Tyr Gln Gly Arg Asp Lys Ser Ile Val Leu Val Ser Phe Val Arg
            740                 745                 750

Ser Asn Lys Asp Gly Thr Val Gly Glu Leu Leu Lys Asp Trp Arg Arg
        755                 760                 765

Leu Asn Val Ala Ile Thr Arg Ala Lys His Lys Leu Ile Leu Leu Gly
    770                 775                 780

Cys Val Pro Ser Leu Asn Cys Tyr Pro Pro Leu Glu Lys Leu Leu Asn
785                 790                 795                 800

His Leu Asn Ser Glu Lys Leu Ile Ile Asp Leu Pro Ser Arg Glu His
                805                 810                 815

Glu Ser Leu Cys His Ile Leu Gly Asp Phe Gln Arg Glu
            820                 825
```

<210> 65

<211> 1522
<212> PRT
<213> Saccharomyces cerevisiae

<220>
<223> DNA2YEAST P38859.1

<400> 65

```
Met Pro Gly Thr Pro Gln Lys Asn Lys Arg Ser Ala Ser Ile Ser Val
1               5                   10                  15

Ser Pro Ala Lys Lys Thr Glu Glu Lys Glu Ile Ile Gln Asn Asp Ser
            20                  25                  30

Lys Ala Ile Leu Ser Lys Gln Thr Lys Arg Lys Lys Lys Tyr Ala Phe
        35                  40                  45

Ala Pro Ile Asn Asn Leu Asn Gly Lys Asn Thr Lys Val Ser Asn Ala
    50                  55                  60

Ser Val Leu Lys Ser Ile Ala Val Ser Gln Val Arg Asn Thr Ser Arg
65                  70                  75                  80

Thr Lys Asp Ile Asn Lys Ala Val Ser Lys Ser Val Lys Gln Leu Pro
                85                  90                  95

Asn Ser Gln Val Lys Pro Lys Arg Glu Met Ser Asn Leu Ser Arg His
            100                 105                 110

His Asp Phe Thr Gln Asp Glu Asp Gly Pro Met Glu Glu Val Ile Trp
```

115 120 125

Lys Tyr Ser Pro Leu Gln Arg Asp Met Ser Asp Lys Thr Thr Ser Ala
130 135 140

Ala Glu Tyr Ser Asp Asp Tyr Glu Asp Val Gln Asn Pro Ser Ser Thr
145 150 155 160

Pro Ile Val Pro Asn Arg Leu Lys Thr Val Leu Ser Phe Thr Asn Ile
165 170 175

Gln Val Pro Asn Ala Asp Val Asn Gln Leu Ile Gln Glu Asn Gly Asn
180 185 190

Glu Gln Val Arg Pro Lys Pro Ala Glu Ile Ser Thr Arg Glu Ser Leu
195 200 205

Arg Asn Ile Asp Asp Ile Leu Asp Asp Ile Glu Gly Asp Leu Thr Ile
210 215 220

Lys Pro Thr Ile Thr Lys Phe Ser Asp Leu Pro Ser Ser Pro Ile Lys
225 230 235 240

Ala Pro Asn Val Glu Lys Lys Ala Glu Val Asn Ala Glu Glu Val Asp
245 250 255

Lys Met Asp Ser Thr Gly Asp Ser Asn Asp Gly Asp Asp Ser Leu Ile
260 265 270

Asp Ile Leu Thr Gln Lys Tyr Val Glu Lys Arg Lys Ser Glu Ser Gln
275 280 285

Ile Thr Ile Gln Gly Asn Thr Asn Gln Lys Ser Gly Ala Gln Glu Ser
290 295 300

Cys Gly Lys Asn Asp Asn Thr Lys Ser Arg Gly Glu Ile Glu Asp His
305 310 315 320

Glu Asn Val Asp Asn Gln Ala Lys Thr Gly Asn Ala Phe Tyr Glu Asn
325 330 335

Glu Glu Asp Ser Asn Cys Gln Arg Ile Lys Lys Asn Glu Lys Ile Glu
340 345 350

Tyr Asn Ser Ser Asp Glu Phe Ser Asp Asp Ser Leu Ile Glu Leu Leu
355 360 365

Asn Glu Thr Gln Thr Gln Val Glu Pro Asn Thr Ile Glu Gln Asp Leu
370 375 380

Asp Lys Val Glu Lys Met Val Ser Asp Asp Leu Arg Ile Ala Thr Asp

385 390 395 400

Ser Thr Leu Ser Ala Tyr Ala Leu Arg Ala Lys Ser Gly Ala Pro Arg
405 410 415

Asp Gly Val Val Arg Leu Val Ile Val Ser Leu Arg Ser Val Glu Leu
420 425 430

Pro Lys Ile Gly Thr Gln Lys Ile Leu Glu Cys Ile Asp Gly Lys Gly
435 440 445

Glu Gln Ser Ser Val Val Val Arg His Pro Trp Val Tyr Leu Glu Phe
450 455 460

Glu Val Gly Asp Val Ile His Ile Ile Glu Gly Lys Asn Ile Glu Asn
465 470 475 480

Lys Arg Leu Leu Ser Asp Asp Lys Asn Pro Lys Thr Gln Leu Ala Asn
485 490 495

Asp Asn Leu Leu Val Leu Asn Pro Asp Val Leu Phe Ser Ala Thr Ser
500 505 510

Val Gly Ser Ser Val Gly Cys Leu Arg Arg Ser Ile Leu Gln Met Gln
515 520 525

Phe Gln Asp Pro Arg Gly Glu Pro Ser Leu Val Met Thr Leu Gly Asn
530 535 540

Ile Val His Glu Leu Leu Gln Asp Ser Ile Lys Tyr Lys Leu Ser His
545 550 555 560

Asn Lys Ile Ser Met Glu Ile Ile Ile Gln Lys Leu Asp Ser Leu Leu
565 570 575

Glu Thr Tyr Ser Phe Ser Ile Ile Ile Cys Asn Glu Glu Ile Gln Tyr
580 585 590

Val Lys Glu Leu Val Met Lys Glu His Ala Glu Asn Ile Leu Tyr Phe
595 600 605

Val Asn Lys Phe Val Ser Lys Ser Asn Tyr Gly Cys Tyr Thr Ser Ile
610 615 620

Ser Gly Thr Arg Arg Thr Gln Pro Ile Ser Ile Ser Asn Val Ile Asp
625 630 635 640

Ile Glu Glu Asn Ile Trp Ser Pro Ile Tyr Gly Leu Lys Gly Phe Leu
645 650 655

Asp Ala Thr Val Glu Ala Asn Val Glu Asn Asn Lys Lys His Ile Val

660 665 670

Pro Leu Glu Val Lys Thr Gly Lys Ser Arg Ser Val Ser Tyr Glu Val
675 680 685

Gln Gly Leu Ile Tyr Thr Leu Leu Leu Asn Asp Arg Tyr Glu Ile Pro
690 695 700

Ile Glu Phe Phe Leu Leu Tyr Phe Thr Arg Asp Lys Asn Met Thr Lys
705 710 715 720

Phe Pro Ser Val Leu His Ser Ile Lys His Ile Leu Met Ser Arg Asn
725 730 735

Arg Met Ser Met Asn Phe Lys His Gln Leu Gln Glu Val Phe Gly Gln
740 745 750

Ala Gln Ser Arg Phe Glu Leu Pro Pro Leu Leu Arg Asp Ser Ser Cys
755 760 765

Asp Ser Cys Phe Ile Lys Glu Ser Cys Met Val Leu Asn Lys Leu Leu
770 775 780

Glu Asp Gly Thr Pro Glu Glu Ser Gly Leu Val Glu Gly Glu Phe Glu
785 790 795 800

Ile Leu Thr Asn His Leu Ser Gln Asn Leu Ala Asn Tyr Lys Glu Phe
805 810 815

Phe Thr Lys Tyr Asn Asp Leu Ile Thr Lys Glu Glu Ser Ser Ile Thr
820 825 830

Cys Val Asn Lys Glu Leu Phe Leu Leu Asp Gly Ser Thr Arg Glu Ser
835 840 845

Arg Ser Gly Arg Cys Leu Ser Gly Leu Val Val Ser Glu Val Val Glu
850 855 860

His Glu Lys Thr Glu Gly Ala Tyr Ile Tyr Cys Phe Ser Arg Arg Arg
865 870 875 880

Asn Asp Asn Asn Ser Gln Ser Met Leu Ser Ser Gln Ile Ala Ala Asn
885 890 895

Asp Phe Val Ile Ile Ser Asp Glu Glu Gly His Phe Cys Leu Cys Gln
900 905 910

Gly Arg Val Gln Phe Ile Asn Pro Ala Lys Ile Gly Ile Ser Val Lys
915 920 925

Arg Lys Leu Leu Asn Asn Arg Leu Leu Asp Lys Glu Lys Gly Val Thr

930 935 940

Thr Ile Gln Ser Val Val Glu Ser Glu Leu Glu Gln Ser Ser Leu Ile
945 950 955 960

Ala Thr Gln Asn Leu Val Thr Tyr Arg Ile Asp Lys Asn Asp Ile Gln
965 970 975

Gln Ser Leu Ser Leu Ala Arg Phe Asn Leu Leu Ser Leu Phe Leu Pro
980 985 990

Ala Val Ser Pro Gly Val Asp Ile Val Asp Glu Arg Ser Lys Leu Cys
995 1000 1005

Arg Lys Thr Lys Arg Ser Asp Gly Gly Asn Glu Ile Leu Arg Ser Leu
1010 1015 1020

Leu Val Asp Asn Arg Ala Pro Lys Phe Arg Asp Ala Asn Asp Asp Pro
1025 1030 1035 1040

Val Ile Pro Tyr Lys Leu Ser Lys Asp Thr Thr Leu Asn Leu Asn Gln
1045 1050 1055

Lys Glu Ala Ile Asp Lys Val Met Arg Ala Glu Asp Tyr Ala Leu Ile
1060 1065 1070

Leu Gly Met Pro Gly Thr Gly Lys Thr Thr Val Ile Ala Glu Ile Ile
1075 1080 1085

Lys Ile Leu Val Ser Glu Gly Lys Arg Val Leu Leu Thr Ser Tyr Thr
1090 1095 1100

His Ser Ala Val Asp Asn Ile Leu Ile Lys Leu Arg Asn Thr Asn Ile
1105 1110 1115 1120

Ser Ile Met Arg Leu Gly Met Lys His Lys Val His Pro Asp Thr Gln
1125 1130 1135

Lys Tyr Val Pro Asn Tyr Ala Ser Val Lys Ser Tyr Asn Asp Tyr Leu
1140 1145 1150

Ser Lys Ile Asn Ser Thr Ser Val Val Ala Thr Thr Cys Leu Gly Ile
1155 1160 1165

Asn Asp Ile Leu Phe Thr Leu Asn Glu Lys Asp Phe Asp Tyr Val Ile
1170 1175 1180

Leu Asp Glu Ala Ser Gln Ile Ser Met Pro Val Ala Leu Gly Pro Leu
1185 1190 1195 1200

Arg Tyr Gly Asn Arg Phe Ile Met Val Gly Asp His Tyr Gln Leu Pro

1205 1210 1215

Pro Leu Val Lys Asn Asp Ala Ala Arg Leu Gly Gly Leu Glu Glu Ser
1220 1225 1230

Leu Phe Lys Thr Phe Cys Glu Lys His Pro Glu Ser Val Ala Glu Leu
1235 1240 1245

Thr Leu Gln Tyr Arg Met Cys Gly Asp Ile Val Thr Leu Ser Asn Phe
1250 1255 1260

Leu Ile Tyr Asp Asn Lys Leu Lys Cys Gly Asn Asn Glu Val Phe Ala
1265 1270 1275 1280

Gln Ser Leu Glu Leu Pro Met Pro Glu Ala Leu Ser Arg Tyr Arg Asn
1285 1290 1295

Glu Ser Ala Asn Ser Lys Gln Trp Leu Glu Asp Ile Leu Glu Pro Thr
1300 1305 1310

Arg Lys Val Val Phe Leu Asn Tyr Asp Asn Cys Pro Asp Ile Ile Glu
1315 1320 1325

Gln Ser Glu Lys Asp Asn Ile Thr Asn His Gly Glu Ala Glu Leu Thr
1330 1335 1340

Leu Gln Cys Val Glu Gly Met Leu Leu Ser Gly Val Pro Cys Glu Asp
1345 1350 1355 1360

Ile Gly Val Met Thr Leu Tyr Arg Ala Gln Leu Arg Leu Leu Lys Lys
1365 1370 1375

Ile Phe Asn Lys Asn Val Tyr Asp Gly Leu Glu Ile Leu Thr Ala Asp
1380 1385 1390

Gln Phe Gln Gly Arg Asp Lys Lys Cys Ile Ile Ile Ser Met Val Arg
1395 1400 1405

Arg Asn Ser Gln Leu Asn Gly Gly Ala Leu Leu Lys Glu Leu Arg Arg
1410 1415 1420

Val Asn Val Ala Met Thr Arg Ala Lys Ser Lys Leu Ile Ile Ile Gly
1425 1430 1435 1440

Ser Lys Ser Thr Ile Gly Ser Val Pro Glu Ile Lys Ser Phe Val Asn
1445 1450 1455

Leu Leu Glu Glu Arg Asn Trp Val Tyr Thr Met Cys Lys Asp Ala Leu
1460 1465 1470

Tyr Lys Tyr Lys Phe Pro Asp Arg Ser Asn Ala Ile Asp Glu Ala Arg

1475 1480 1485

Lys Gly Cys Gly Lys Arg Thr Gly Ala Lys Pro Ile Thr Ser Lys Ser
1490 1495 1500

Lys Phe Val Ser Asp Lys Pro Ile Ile Lys Glu Ile Leu Gln Glu Tyr
1505 1510 1515 1520

Glu Ser

<210> 66

<211> 490
<212> PRT
<213> Human herpesvirus 2

<220>
<223> VP16 AAA45863.1

<400> 66
Met Asp Leu Leu Val Asp Asp Leu Phe Ala Asp Arg Asp Gly Val Ser
1 5 10 15

Pro Pro Pro Pro Arg Pro Ala Gly Gly Pro Lys Asn Thr Pro Ala Ala
20 25 30

Pro Pro Leu Tyr Ala Thr Gly Arg Leu Ser Gln Ala Gln Leu Met Pro
35 40 45

Ser Pro Pro Met Pro Val Pro Pro Ala Ala Leu Phe Asn Arg Leu Leu
50 55 60

Asp Asp Leu Gly Phe Ser Ala Gly Pro Ala Leu Cys Thr Met Leu Asp
65 70 75 80

Thr Trp Asn Glu Asp Leu Phe Ser Gly Phe Pro Thr Asn Ala Asp Met
85 90 95

Tyr Arg Glu Cys Lys Phe Leu Ser Thr Leu Pro Ser Asp Val Ile Asp
100 105 110

Trp Gly Asp Ala His Val Pro Glu Arg Ser Pro Ile Asp Ile Arg Ala
115 120 125

His Gly Asp Val Ala Phe Pro Thr Leu Pro Ala Thr Arg Asp Glu Leu
130 135 140

Pro Ser Tyr Tyr Glu Ala Met Ala Gln Phe Phe Arg Gly Glu Leu Arg
145 150 155 160

Ala Arg Glu Glu Ser Tyr Arg Thr Val Leu Ala Asn Phe Cys Ser Ala
165 170 175

```
Leu Tyr Arg Tyr Leu Arg Ala Ser Val Arg Gln Leu His Arg Gln Ala
            180                 185                 190

His Met Arg Gly Arg Asn Arg Asp Leu Arg Glu Met Leu Arg Thr Thr
        195                 200                 205

Ile Ala Asp Arg Tyr Tyr Arg Glu Thr Ala Arg Leu Ala Arg Val Leu
    210                 215                 220

Phe Leu His Leu Tyr Leu Phe Leu Ser Arg Glu Ile Leu Trp Ala Ala
225                 230                 235                 240

Tyr Ala Glu Gln Met Met Arg Pro Asp Leu Phe Asp Gly Leu Cys Cys
                245                 250                 255

Asp Leu Glu Ser Trp Arg Gln Leu Ala Cys Leu Phe Gln Pro Leu Met
            260                 265                 270

Phe Ile Asn Gly Ser Leu Thr Val Arg Gly Val Pro Val Glu Ala Arg
        275                 280                 285

Arg Leu Arg Glu Leu Asn His Ile Arg Glu His Leu Asn Leu Pro Leu
    290                 295                 300

Val Arg Ser Ala Ala Ala Glu Glu Pro Gly Ala Pro Leu Thr Thr Pro
305                 310                 315                 320

Pro Val Leu Gln Gly Asn Gln Ala Arg Ser Ser Gly Tyr Phe Met Leu
                325                 330                 335

Leu Ile Arg Ala Lys Leu Asp Ser Tyr Ser Ser Val Ala Thr Ser Glu
            340                 345                 350

Gly Glu Ser Val Met Arg Glu His Ala Tyr Ser Arg Gly Arg Thr Arg
        355                 360                 365

Asn Asn Tyr Gly Ser Thr Ile Glu Gly Leu Leu Asp Leu Pro Asp Asp
    370                 375                 380

Asp Asp Ala Pro Ala Glu Ala Gly Leu Val Ala Pro Arg Met Ser Phe
385                 390                 395                 400

Leu Ser Ala Gly Gln Arg Pro Arg Arg Leu Ser Thr Thr Ala Pro Ile
                405                 410                 415

Thr Asp Val Ser Leu Gly Asp Glu Leu Arg Leu Asp Gly Glu Glu Val
            420                 425                 430

Asp Met Thr Pro Ala Asp Ala Leu Asp Asp Phe Asp Leu Glu Met Leu
        435                 440                 445
```

Gly Asp Val Glu Ser Pro Ser Pro Gly Met Thr His Asp Pro Val Ser
450 455 460

Tyr Gly Ala Leu Asp Val Asp Asp Phe Glu Phe Glu Gln Met Phe Thr
465 470 475 480

Asp Ala Met Gly Ile Asp Asp Phe Gly Gly
485 490

<210> 67

<211> 3
<212> PRT

<220>
<223> 1a8h_1 peptidic linker

<400> 67
Asn Val Gly
1

<210> 68

<211> 4
<212> PRT

<220>
<223> 1dnpA_1 peptidic linker

<400> 68
Asp Ser Val Ile
1

<210> 69

<211> 4
<212> PRT

<220>
<223> 1d8cA_2 peptidic linker

<400> 69
Ile Val Glu Ala
1

<210> 70

<211> 4
<212> PRT

<220>
<223> 1ckqA_3 peptidic linker

<400> 70
Leu Glu Gly Ser
1

<210> 71

<211> 4
<212> PRT

<220>
<223> 1sbp_1 peptidic linker

<400> 71
Tyr Thr Ser Thr
1

<210> 72

<211> 5
<212> PRT

<220>
<223> 1ev7A_1 peptidic linker

<400> 72
Leu Gln Glu Asn Leu
1   5

<210> 73

<211> 5
<212> PRT

<220>
<223> 1alo_3 peptidic linker

<400> 73
Val Gly Arg Gln Pro
1   5

<210> 74

<211> 5
<212> PRT

<220>
<223> 1amf_1 peptidic linker

<400> 74
Leu Gly Asn Ser Leu
1   5

<210> 75

<211> 6
<212> PRT

<220>
<223> 1adjA_3 peptidic linker

<400> 75
Leu Pro Glu Glu Lys Gly
1   5

<210> 76

<211> 6
<212> PRT

<220>
<223> 1fcdC_1 peptidic linker

<400> 76
Gln Thr Tyr Gln Pro Ala
1 5

<210> 77

<211> 6
<212> PRT

<220>
<223> 1al3_2 peptidic linker

<400> 77
Phe Ser His Ser Thr Thr
1 5

<210> 78

<211> 7
<212> PRT

<220>
<223> 1g3p_1 peptidic linker

<400> 78
Gly Tyr Thr Tyr Ile Asn Pro
1 5

<210> 79

<211> 7
<212> PRT

<220>
<223> 1acc_3 peptidic linker

<400> 79
Leu Thr Lys Tyr Lys Ser Ser
1 5

<210> 80

<211> 8
<212> PRT

<220>
<223> 1ahjB_1 peptidic linker

<400> 80
Ser Arg Pro Ser Glu Ser Glu Gly
1 5

<210> 81

<211> 8
<212> PRT

<220>
<223> 1acc_1 peptidic linker

<400> 81
Pro Glu Leu Lys Gln Lys Ser Ser

1 5

<210> 82

<211> 8
<212> PRT

<220>
<223> 1af7_1 peptidic linker

<400> 82
Leu Thr Thr Asn Leu Thr Ala Phe
1 5

<210> 83

<211> 9
<212> PRT

<220>
<223> 1heiA_1 peptidic linker

<400> 83
Thr Ala Thr Pro Pro Gly Ser Val Thr
1 5

<210> 84

<211> 9
<212> PRT

<220>
<223> 1bia_2 peptidic linker

<400> 84
Leu Asp Asn Phe Ile Asn Arg Pro Val
1 5

<210> 85

<211> 9
<212> PRT

<220>
<223> 1igtB_1 peptidic linker

<400> 85
Val Ser Ser Ala Lys Thr Thr Ala Pro
1 5

<210> 86

<211> 10
<212> PRT

<220>
<223> 1nfkA_1 peptidic linker

<400> 86
Asp Ser Lys Ala Pro Asn Ala Ser Asn Leu
1 5 10

<210> 87

<211> 10
<212> PRT

<220>
<223> 1au7A_1 peptidic linker

<400> 87
Lys Arg Arg Thr Thr Ile Ser Ile Ala Ala
1 5 10

<210> 88

<211> 11
<212> PRT

<220>
<223> 1bpoB_1 peptidic linker

<400> 88
Pro Val Lys Met Phe Asp Arg His Ser Ser Leu
1 5 10

<210> 89

<211> 11
<212> PRT

<220>
<223> 1b0pA_2 peptidic linker

<400> 89
Ala Pro Ala Glu Thr Lys Ala Glu Pro Met Thr
1 5 10

<210> 90

<211> 14
<212> PRT

<220>
<223> 1c05A_2 peptidic linker

<400> 90
Tyr Thr Arg Leu Pro Glu Arg Ser Glu Leu Pro Ala Glu Ile
1 5 10

<210> 91

<211> 14
<212> PRT

<220>
<223> 1gcb_1 peptidic linker

<400> 91
Val Ser Thr Asp Ser Thr Pro Val Thr Asn Gln Lys Ser Ser
1 5 10

<210> 92

<211> 14

<212> PRT

<220>
<223> 1bt3A_1 peptidic linker

<400> 92
Tyr Lys Leu Pro Ala Val Thr Thr Met Lys Val Arg Pro Ala
1   5   10

<210> 93

<211> 15
<212> PRT

<220>
<223> 1b3oB_2 peptidic linker

<400> 93
Ile Ala Arg Thr Asp Leu Lys Lys Asn Arg Asp Tyr Pro Leu Ala
1   5   10   15

<210> 94

<211> 21
<212> PRT

<220>
<223> 16vpA_6 peptidic linker

<400> 94
Thr Glu Glu Pro Gly Ala Pro Leu Thr Thr Pro Pro Thr Leu His Gly
1   5   10   15

Asn Gln Ala Arg Ala
            20

<210> 95

<211> 21
<212> PRT

<220>
<223> 1dhx_1 peptidic linker

<400> 95
Ala Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met Ala
1   5   10   15

Ser Thr Tyr Phe Asp
            20

<210> 96

<211> 26
<212> PRT

<220>
<223> 1b8aA_1 peptidic linker

<400> 96
Ile Val Val Leu Asn Arg Ala Glu Thr Pro Leu Pro Leu Asp Pro Thr
1   5   10   15

Gly Lys Val Lys Ala Glu Leu Asp Thr Arg
20 25

<210> 97

<211> 28
<212> PRT

<220>
<223> 1qu6A_1 peptidic linker

<400> 97
Ile Leu Asn Lys Glu Lys Lys Ala Val Ser Pro Leu Leu Leu Thr Thr
1 5 10 15

Thr Asn Ser Ser Glu Gly Leu Ser Met Gly Asn Tyr
20 25

<210> 98

<211> 20
<212> PRT

<220>
<223> NFS1 peptidic linker

<400> 98
Gly Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu Lys Met Lys Gly Gln
1 5 10 15

Gly Pro Ser Gly
20

<210> 99

<211> 23
<212> PRT

<220>
<223> NFS2 peptidic linker

<400> 99
Gly Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu Lys Met Lys Gly Leu
1 5 10 15

Gly Pro Asp Gly Arg Lys Ala
20

<210> 100

<211> 10
<212> PRT

<220>
<223> CFS1 peptidic linker

<400> 100
Ser Leu Thr Lys Ser Lys Ile Ser Gly Ser
1 5 10

<210> 101

<211> 32
<212> PRT

<220>
<223> RM2 peptidic linker

<400> 101
Ala Ala Gly Gly Ser Ala Leu Thr Ala Gly Ala Leu Ser Leu Thr Ala
1 5 10 15

Gly Ala Leu Ser Leu Thr Ala Gly Ala Leu Ser Gly Gly Gly Gly Ser
20 25 30

<210> 102

<211> 27
<212> PRT

<220>
<223> BQY peptidic linker

<400> 102
Ala Ala Gly Ala Ser Ser Val Ser Ala Ser Gly His Ile Ala Pro Leu
1 5 10 15

Ser Leu Pro Ser Ser Pro Pro Ser Val Gly Ser
20 25

<210> 103

<211> 5
<212> PRT

<220>
<223> QGPSG peptidic linker

<400> 103
Gln Gly Pro Ser Gly
1 5

<210> 104

<211> 8
<212> PRT

<220>
<223> LGPDGRKA peptidic linker

<400> 104
Leu Gly Pro Asp Gly Arg Lys Ala
1 5

<210> 105
<211> 43
<212> DNA
<213> Artificial

<220>

<223> I-TevI target

<400> 105
caacgctcag tagatgtttt cttgggtcta ccgtttaata ttg 43

<210> 106

<211> 167
<212> PRT
<213> artificial sequence

<220>
<223> I-CreI

<400> 106
Met Ala Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly
1 5 10 15

Phe Val Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln
20 25 30

Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln
35 40 45

Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
50 55 60

Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser
65 70 75 80

Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu
85 90 95

Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln
100 105 110

Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr
115 120 125

Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr
130 135 140

Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys
145 150 155 160

Lys Ser Ser Pro Ala Ala Asp
165

<210> 107

<211> 245
<212> PRT
<213> artificial sequence

<220>
<223> I-TevI

<400> 107

```
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1               5                  10                  15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
            20                  25                  30

Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
        35                  40                  45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
    50                  55                  60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65                  70                  75                  80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
                85                  90                  95

Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
            100                 105                 110

Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly
        115                 120                 125

Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp Asn
    130                 135                 140

Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr Ser
145                 150                 155                 160

Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn Ser
                165                 170                 175

Phe Phe Asn His Lys His Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu
            180                 185                 190

Lys Met Lys Gly Lys Lys Pro Ser Asn Ile Lys Lys Ile Ser Cys Asp
        195                 200                 205

Gly Val Ile Phe Asp Cys Ala Ala Asp Ala Ala Arg His Phe Lys Ile
    210                 215                 220

Ser Ser Gly Leu Val Thr Tyr Arg Val Lys Ser Asp Lys Trp Asn Trp
225                 230                 235                 240

Phe Tyr Ile Asn Ala
                245
```

<210> 108

<211> 167

<212> PRT
<213> artificial sequence

<220>
<223> I-CreI_N20

<400> 108
Met Ala Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly
1 5 10 15

Phe Val Asp Gly Asn Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln
20 25 30

Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln
35 40 45

Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
50 55 60

Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser
65 70 75 80

Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu
85 90 95

Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln
100 105 110

Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr
115 120 125

Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr
130 135 140

Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys
145 150 155 160

Lys Ser Ser Pro Ala Ala Asp
165

<210> 109

<211> 366
<212> PRT
<213> artificial sequence

<220>
<223> hTevCre_D01

<400> 109
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1 5 10 15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
20 25 30

```
Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
        35                  40                  45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
    50                  55                  60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65                  70                  75                  80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
                85                  90                  95

Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
            100                 105                 110

Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly
        115                 120                 125

Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp Asn
    130                 135                 140

Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr Ser
145                 150                 155                 160

Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn Ser
                165                 170                 175

Phe Phe Asn His Lys His Ser Asp Ile Thr Lys Ser Lys Ile Ser Gly
            180                 185                 190

Lys Met Lys Gly Gln Gly Pro Ser Gly Asn Thr Lys Tyr Asn Lys Glu
        195                 200                 205

Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile Ile
    210                 215                 220

Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Gln Leu Ser
225                 230                 235                 240

Leu Thr Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp
                245                 250                 255

Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp Arg Gly Ser
            260                 265                 270

Val Ser Asp Tyr Ile Leu Ser Glu Ile Lys Pro Leu His Asn Phe Leu
        275                 280                 285

Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn Leu
    290                 295                 300
```

Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp
305 310 315 320

Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu Asn
325 330 335

Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val Leu
340 345 350

Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro Ala Ala Asp
355 360 365

<210> 110

<211> 353
<212> PRT
<213> artificial sequence

<220>
<223> hTevCre_D02

<400> 110
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1 5 10 15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
20 25 30

Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
35 40 45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
50 55 60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65 70 75 80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
85 90 95

Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
100 105 110

Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly
115 120 125

Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp Asn
130 135 140

Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr Ser
145 150 155 160

Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn Ser

165 170 175

Phe Phe Asn His Lys His Ser Gln Gly Pro Ser Gly Asn Thr Lys Tyr
180 185 190

Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly
195 200 205

Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His
210 215 220

Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp
225 230 235 240

Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp
245 250 255

Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu Ile Lys Pro Leu His
260 265 270

Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln
275 280 285

Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu
290 295 300

Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala
305 310 315 320

Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg
325 330 335

Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro Ala Ala
340 345 350

Asp

<210> 111

<211> 339
<212> PRT
<213> artificial sequence

<220>
<223> hTevCre_D03

<400> 111
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1 5 10 15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
20 25 30

```
Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
        35                  40                  45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
    50                  55                  60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65                  70                  75                  80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
                85                  90                  95

Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
            100                 105                 110

Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly
        115                 120                 125

Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp Asn
    130                 135                 140

Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr Ser
145                 150                 155                 160

Ala Tyr Thr Cys Ser Lys Cys Arg Asn Gln Gly Pro Ser Gly Asn Thr
                165                 170                 175

Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly
            180                 185                 190

Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe
        195                 200                 205

Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys Thr Gln Arg
    210                 215                 220

Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val
225                 230                 235                 240

Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu Ile Lys Pro
                245                 250                 255

Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln
            260                 265                 270

Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala
        275                 280                 285

Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln
    290                 295                 300
```

Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr
305 310 315 320

Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro
325 330 335

Ala Ala Asp

<210> 112

<211> 318
<212> PRT
<213> artificial sequence

<220>
<223> hTevCre_D04

<400> 112
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1 5 10 15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
20 25 30

Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
35 40 45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
50 55 60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65 70 75 80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
85 90 95

Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
100 105 110

Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly
115 120 125

Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp Asn
130 135 140

Pro Glu Thr His Gln Gly Pro Ser Gly Asn Thr Lys Tyr Asn Lys Glu
145 150 155 160

Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile Ile
165 170 175

Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Gln Leu Ser

180 185 190

Leu Thr Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp
195 200 205

Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp Arg Gly Ser
210 215 220

Val Ser Asp Tyr Ile Leu Ser Glu Ile Lys Pro Leu His Asn Phe Leu
225 230 235 240

Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn Leu
245 250 255

Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp
260 265 270

Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu Asn
275 280 285

Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val Leu
290 295 300

Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro Ala Ala Asp
305 310 315

<210> 113

<211> 293
<212> PRT
<213> artificial sequence

<220>
<223> hTevCre_D05

<400> 113
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1 5 10 15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
20 25 30

Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
35 40 45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
50 55 60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65 70 75 80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
85 90 95

```
Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
            100                 105                 110

Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Gln Gly Pro Ser Gly
        115                 120                 125

Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val
    130                 135                 140

Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr
145                 150                 155                 160

Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys Thr
                165                 170                 175

Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly
            180                 185                 190

Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu Ile
        195                 200                 205

Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu
    210                 215                 220

Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro
225                 230                 235                 240

Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val
                245                 250                 255

Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser
            260                 265                 270

Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser
        275                 280                 285

Ser Pro Ala Ala Asp
    290


<210>  114

<211>  296
<212>  PRT
<213>  artificial sequence

<220>
<223>  hTevCre_D06

<400>  114
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1               5                   10                  15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
```

20 25 30

Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
35 40 45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
50 55 60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65 70 75 80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
85 90 95

Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
100 105 110

Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly
115 120 125

Arg Lys Ala Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala
130 135 140

Gly Phe Val Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn
145 150 155 160

Gln Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr
165 170 175

Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile
180 185 190

Gly Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu
195 200 205

Ser Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe
210 215 220

Leu Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu
225 230 235 240

Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys
245 250 255

Thr Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys
260 265 270

Thr Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys
275 280 285

Lys Lys Ser Ser Pro Ala Ala Asp

290 295

<210> 115

<211> 366
<212> PRT
<213> artificial sequence

<220>
<223> hTevCre_D01_N20

<400> 115
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1 5 10 15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
20 25 30

Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
35 40 45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
50 55 60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65 70 75 80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
85 90 95

Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
100 105 110

Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly
115 120 125

Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp Asn
130 135 140

Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr Ser
145 150 155 160

Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn Ser
165 170 175

Phe Phe Asn His Lys His Ser Asp Ile Thr Lys Ser Lys Ile Ser Gly
180 185 190

Lys Met Lys Gly Gln Gly Pro Ser Gly Asn Thr Lys Tyr Asn Lys Glu
195 200 205

Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asn Gly Ser Ile Ile
210 215 220

Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Gln Leu Ser
225 230 235 240

Leu Thr Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp
245 250 255

Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp Arg Gly Ser
260 265 270

Val Ser Asp Tyr Ile Leu Ser Glu Ile Lys Pro Leu His Asn Phe Leu
275 280 285

Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn Leu
290 295 300

Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp
305 310 315 320

Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu Asn
325 330 335

Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val Leu
340 345 350

Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro Ala Ala Asp
355 360 365

<210> 116

<211> 353
<212> PRT
<213> artificial sequence

<220>
<223> hTevCre_D02_N20

<400> 116
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1 5 10 15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
20 25 30

Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
35 40 45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
50 55 60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65 70 75 80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe

85 90 95

Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
100 105 110

Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly
115 120 125

Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp Asn
130 135 140

Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr Ser
145 150 155 160

Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn Ser
165 170 175

Phe Phe Asn His Lys His Ser Gln Gly Pro Ser Gly Asn Thr Lys Tyr
180 185 190

Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asn Gly
195 200 205

Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His
210 215 220

Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp
225 230 235 240

Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp
245 250 255

Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu Ile Lys Pro Leu His
260 265 270

Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln
275 280 285

Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu
290 295 300

Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala
305 310 315 320

Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg
325 330 335

Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro Ala Ala
340 345 350

Asp

<210> 117

<211> 339
<212> PRT
<213> artificial sequence

<220>
<223> hTevCre_D03_N20

<400> 117
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1 5 10 15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
20 25 30

Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
35 40 45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
50 55 60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65 70 75 80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
85 90 95

Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
100 105 110

Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly
115 120 125

Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp Asn
130 135 140

Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr Ser
145 150 155 160

Ala Tyr Thr Cys Ser Lys Cys Arg Asn Gln Gly Pro Ser Gly Asn Thr
165 170 175

Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly
180 185 190

Asn Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe
195 200 205

Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys Thr Gln Arg
210 215 220

Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val
225 230 235 240

Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu Ile Lys Pro
245 250 255

Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln
260 265 270

Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala
275 280 285

Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln
290 295 300

Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr
305 310 315 320

Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro
325 330 335

Ala Ala Asp

<210> 118

<211> 318
<212> PRT
<213> artificial sequence

<220>
<223> hTevCre_D04_N20

<400> 118
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1 5 10 15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
20 25 30

Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
35 40 45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
50 55 60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65 70 75 80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
85 90 95

Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
100 105 110

```
Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly
        115                 120                 125

Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp Asn
    130                 135                 140

Pro Glu Thr His Gln Gly Pro Ser Gly Asn Thr Lys Tyr Asn Lys Glu
145                 150                 155                 160

Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asn Gly Ser Ile Ile
                165                 170                 175

Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Gln Leu Ser
            180                 185                 190

Leu Thr Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp
        195                 200                 205

Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp Arg Gly Ser
    210                 215                 220

Val Ser Asp Tyr Ile Leu Ser Glu Ile Lys Pro Leu His Asn Phe Leu
225                 230                 235                 240

Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn Leu
                245                 250                 255

Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp
            260                 265                 270

Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu Asn
        275                 280                 285

Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val Leu
    290                 295                 300

Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro Ala Ala Asp
305                 310                 315


<210>  119

<211>  293
<212>  PRT
<213>  artificial sequence

<220>
<223>  hTevCre_D05_N20

<400>  119
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1               5                   10                  15
```

```
Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
            20                  25                  30

Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
        35                  40                  45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
    50                  55                  60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65                  70                  75                  80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
                85                  90                  95

Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
            100                 105                 110

Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Gln Gly Pro Ser Gly
        115                 120                 125

Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val
    130                 135                 140

Asp Gly Asn Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr
145                 150                 155                 160

Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys Thr
                165                 170                 175

Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly
            180                 185                 190

Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu Ile
        195                 200                 205

Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu
    210                 215                 220

Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro
225                 230                 235                 240

Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val
                245                 250                 255

Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser
            260                 265                 270

Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser
        275                 280                 285
```

Ser Pro Ala Ala Asp
290

<210> 120

<211> 296
<212> PRT
<213> artificial sequence

<220>
<223> hTevCre_D06_N20

<400> 120
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1 5 10 15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
20 25 30

Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
35 40 45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
50 55 60

Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65 70 75 80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
85 90 95

Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys
100 105 110

Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly
115 120 125

Arg Lys Ala Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala
130 135 140

Gly Phe Val Asp Gly Asn Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn
145 150 155 160

Gln Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr
165 170 175

Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile
180 185 190

Gly Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu
195 200 205

Ser Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe
210 215 220

Leu Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu
225 230 235 240

Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys
245 250 255

Thr Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys
260 265 270

Thr Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys
275 280 285

Lys Lys Ser Ser Pro Ala Ala Asp
290 295

<210> 121

<211> 157
<212> PRT
<213> artificial sequence

<220>
<223> I-CreI_X

<400> 121
Met Ala Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly
1 5 10 15

Phe Val Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln
20 25 30

Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln
35 40 45

Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
50 55 60

Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser
65 70 75 80

Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu
85 90 95

Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln
100 105 110

Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr
115 120 125

Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr
130 135 140

Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Ala Asp
145 150 155

<210> 122

<211> 177
<212> PRT
<213> artificial sequence

<220>
<223> I-CreI_NFS1

<400> 122
Met Ala Gly Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu Lys Met Lys
1 5 10 15

Gly Gln Gly Pro Ser Gly Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu
20 25 30

Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile
35 40 45

Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe
50 55 60

Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val
65 70 75 80

Asp Glu Ile Gly Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp
85 90 95

Tyr Ile Leu Ser Glu Ile Ala Pro Leu His Asn Phe Leu Thr Gln Leu
100 105 110

Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys
115 120 125

Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu
130 135 140

Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys
145 150 155 160

Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Ala
165 170 175

Asp

<210> 123

<211> 180
<212> PRT
<213> artificial sequence

<220>

<223> I-CreI_NFS2

<400> 123
Met Ala Gly Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu Lys Met Lys
1 5 10 15

Gly Leu Gly Pro Asp Gly Arg Lys Ala Asn Thr Lys Tyr Asn Lys Glu
20 25 30

Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile Ile
35 40 45

Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Gln Leu Ser
50 55 60

Leu Thr Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp
65 70 75 80

Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp Arg Gly Ser
85 90 95

Val Ser Asp Tyr Ile Leu Ser Glu Ile Ala Pro Leu His Asn Phe Leu
100 105 110

Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn Leu
115 120 125

Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp
130 135 140

Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu Asn
145 150 155 160

Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val Leu
165 170 175

Asp Ser Ala Asp
180

<210> 124

<211> 166
<212> PRT
<213> artificial sequence

<220>
<223> I-CreI_CFS1

<400> 124
Met Ala Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly
1 5 10 15

Phe Val Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln
20 25 30

Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln
35 40 45

Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
50 55 60

Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser
65 70 75 80

Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu
85 90 95

Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln
100 105 110

Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr
115 120 125

Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr
130 135 140

Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Thr Lys Ser Lys
145 150 155 160

Ile Ser Gly Ser Ala Asp
165

<210> 125

<211> 177
<212> PRT
<213> artificial sequence

<220>
<223> I-CreI_NFS1_N20

<400> 125
Met Ala Gly Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu Lys Met Lys
1 5 10 15

Gly Gln Gly Pro Ser Gly Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu
20 25 30

Tyr Leu Ala Gly Phe Val Asp Gly Asn Gly Ser Ile Ile Ala Gln Ile
35 40 45

Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe
50 55 60

Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val
65 70 75 80

Asp Glu Ile Gly Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp

85 90 95

Tyr Ile Leu Ser Glu Ile Ala Pro Leu His Asn Phe Leu Thr Gln Leu
100 105 110

Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys
115 120 125

Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu
130 135 140

Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys
145 150 155 160

Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Ala
165 170 175

Asp

<210> 126

<211> 180
<212> PRT
<213> artificial sequence

<220>
<223> I-CreI_NFS2_N20

<400> 126
Met Ala Gly Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu Lys Met Lys
1 5 10 15

Gly Leu Gly Pro Asp Gly Arg Lys Ala Asn Thr Lys Tyr Asn Lys Glu
20 25 30

Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asn Gly Ser Ile Ile
35 40 45

Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Gln Leu Ser
50 55 60

Leu Thr Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp
65 70 75 80

Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp Arg Gly Ser
85 90 95

Val Ser Asp Tyr Ile Leu Ser Glu Ile Ala Pro Leu His Asn Phe Leu
100 105 110

Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn Leu
115 120 125

Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp
130 135 140

Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu Asn
145 150 155 160

Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val Leu
165 170 175

Asp Ser Ala Asp
180

<210> 127

<211> 166
<212> PRT
<213> artificial sequence

<220>
<223> I-CreI_CFS1_N20

<400> 127
Met Ala Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly
1 5 10 15

Phe Val Asp Gly Asn Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln
20 25 30

Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln
35 40 45

Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
50 55 60

Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser
65 70 75 80

Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu
85 90 95

Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln
100 105 110

Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr
115 120 125

Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr
130 135 140

Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Thr Lys Ser Lys
145 150 155 160

Ile Ser Gly Ser Ala Asp

165

<210> 128

<211> 311
<212> PRT
<213> artificial sequence

<220>
<223> hColE7Cre_D0101

<400> 128

```
Met Ala Ser Gly Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys
1               5                   10                  15

Gly Lys Pro Val Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu
            20                  25                  30

Gly Ser Pro Val Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu
        35                  40                  45

Phe Lys Ser Phe Asp Asp Phe Arg Lys Lys Phe Trp Glu Glu Val Ser
    50                  55                  60

Lys Asp Pro Glu Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg
65                  70                  75                  80

Met Lys Val Gly Lys Ala Pro Lys Thr Arg Thr Gln Asp Val Ser Gly
                85                  90                  95

Lys Arg Thr Ser Phe Glu Leu His His Glu Lys Pro Ile Ser Gln Asn
            100                 105                 110

Gly Gly Val Tyr Asp Met Asp Asn Ile Ser Val Val Thr Pro Lys Arg
        115                 120                 125

His Ile Asp Ile His Arg Gly Lys Gly Ser Asp Ile Thr Lys Ser Lys
    130                 135                 140

Ile Ser Glu Lys Met Lys Gly Gln Gly Pro Ser Gly Asn Thr Lys Tyr
145                 150                 155                 160

Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly
                165                 170                 175

Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His
            180                 185                 190

Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp
        195                 200                 205

Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp
    210                 215                 220
```

Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu Ile Ala Pro Leu His
225 230 235 240

Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln
245 250 255

Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu
260 265 270

Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala
275 280 285

Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg
290 295 300

Ala Val Leu Asp Ser Ala Asp
305 310

<210> 129

<211> 314
<212> PRT
<213> artificial sequence

<220>
<223> hColE7Cre_D0102

<400> 129
Met Ala Ser Gly Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys
1 5 10 15

Gly Lys Pro Val Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu
20 25 30

Gly Ser Pro Val Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu
35 40 45

Phe Lys Ser Phe Asp Asp Phe Arg Lys Lys Phe Trp Glu Glu Val Ser
50 55 60

Lys Asp Pro Glu Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg
65 70 75 80

Met Lys Val Gly Lys Ala Pro Lys Thr Arg Thr Gln Asp Val Ser Gly
85 90 95

Lys Arg Thr Ser Phe Glu Leu His His Glu Lys Pro Ile Ser Gln Asn
100 105 110

Gly Gly Val Tyr Asp Met Asp Asn Ile Ser Val Val Thr Pro Lys Arg
115 120 125

His Ile Asp Ile His Arg Gly Lys Gly Ser Asp Ile Thr Lys Ser Lys

130 135 140

Ile Ser Glu Lys Met Lys Gly Leu Gly Pro Asp Gly Arg Lys Ala Asn
145 150 155 160

Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp
165 170 175

Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys
180 185 190

Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys Thr Gln
195 200 205

Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr
210 215 220

Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu Ile Ala
225 230 235 240

Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys
245 250 255

Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser
260 265 270

Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp
275 280 285

Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu
290 295 300

Thr Val Arg Ala Val Leu Asp Ser Ala Asp
305 310

<210> 130

<211> 300
<212> PRT
<213> artificial sequence

<220>
<223> hCreColE7_D0101

<400> 130
Met Ala Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly
1 5 10 15

Phe Val Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln
20 25 30

Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln
35 40 45

```
Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
    50                  55                  60

Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser
65                  70                  75                  80

Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu
                85                  90                  95

Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln
            100                 105                 110

Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr
        115                 120                 125

Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr
    130                 135                 140

Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Thr Lys Ser Lys
145                 150                 155                 160

Ile Ser Gly Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys Gly
                165                 170                 175

Lys Pro Val Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu Gly
            180                 185                 190

Ser Pro Val Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu Phe
        195                 200                 205

Lys Ser Phe Asp Asp Phe Arg Lys Lys Phe Trp Glu Glu Val Ser Lys
    210                 215                 220

Asp Pro Glu Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg Met
225                 230                 235                 240

Lys Val Gly Lys Ala Pro Lys Thr Arg Thr Gln Asp Val Ser Gly Lys
                245                 250                 255

Arg Thr Ser Phe Glu Leu His His Glu Lys Pro Ile Ser Gln Asn Gly
            260                 265                 270

Gly Val Tyr Asp Met Asp Asn Ile Ser Val Val Thr Pro Lys Arg His
        275                 280                 285

Ile Asp Ile His Arg Gly Lys Gly Ser Ser Ala Asp
    290                 295                 300
```

<210> 131

<211> 311

<212> PRT
<213> artificial sequence

<220>
<223> hColE7Cre_D0101_N20

<400> 131

```
Met Ala Ser Gly Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys
1               5                  10                  15

Gly Lys Pro Val Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu
            20                  25                  30

Gly Ser Pro Val Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu
        35                  40                  45

Phe Lys Ser Phe Asp Asp Phe Arg Lys Lys Phe Trp Glu Glu Val Ser
    50                  55                  60

Lys Asp Pro Glu Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg
65                  70                  75                  80

Met Lys Val Gly Lys Ala Pro Lys Thr Arg Thr Gln Asp Val Ser Gly
                85                  90                  95

Lys Arg Thr Ser Phe Glu Leu His His Glu Lys Pro Ile Ser Gln Asn
            100                 105                 110

Gly Gly Val Tyr Asp Met Asp Asn Ile Ser Val Val Thr Pro Lys Arg
        115                 120                 125

His Ile Asp Ile His Arg Gly Lys Gly Ser Asp Ile Thr Lys Ser Lys
    130                 135                 140

Ile Ser Glu Lys Met Lys Gly Gln Gly Pro Ser Gly Asn Thr Lys Tyr
145                 150                 155                 160

Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asn Gly
                165                 170                 175

Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His
            180                 185                 190

Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp
        195                 200                 205

Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp
    210                 215                 220

Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu Ile Ala Pro Leu His
225                 230                 235                 240

Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln
```

245 250 255

Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu
260 265 270

Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala
275 280 285

Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg
290 295 300

Ala Val Leu Asp Ser Ala Asp
305 310

<210> 132

<211> 314
<212> PRT
<213> artificial sequence

<220>
<223> hColE7Cre_D0102_N20

<400> 132
Met Ala Ser Gly Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys
1 5 10 15

Gly Lys Pro Val Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu
20 25 30

Gly Ser Pro Val Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu
35 40 45

Phe Lys Ser Phe Asp Asp Phe Arg Lys Lys Phe Trp Glu Glu Val Ser
50 55 60

Lys Asp Pro Glu Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg
65 70 75 80

Met Lys Val Gly Lys Ala Pro Lys Thr Arg Thr Gln Asp Val Ser Gly
85 90 95

Lys Arg Thr Ser Phe Glu Leu His His Glu Lys Pro Ile Ser Gln Asn
100 105 110

Gly Gly Val Tyr Asp Met Asp Asn Ile Ser Val Val Thr Pro Lys Arg
115 120 125

His Ile Asp Ile His Arg Gly Lys Gly Ser Asp Ile Thr Lys Ser Lys
130 135 140

Ile Ser Glu Lys Met Lys Gly Leu Gly Pro Asp Gly Arg Lys Ala Asn
145 150 155 160

Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp
165 170 175

Gly Asn Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys
180 185 190

Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys Thr Gln
195 200 205

Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr
210 215 220

Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu Ile Ala
225 230 235 240

Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys
245 250 255

Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser
260 265 270

Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp
275 280 285

Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu
290 295 300

Thr Val Arg Ala Val Leu Asp Ser Ala Asp
305 310

<210> 133

<211> 300
<212> PRT
<213> artificial sequence

<220>
<223> hCreColE7_D0101_N20

<400> 133
Met Ala Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly
1 5 10 15

Phe Val Asp Gly Asn Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln
20 25 30

Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln
35 40 45

Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
50 55 60

Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser

65 70 75 80

Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu
85 90 95

Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln
100 105 110

Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr
115 120 125

Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr
130 135 140

Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Thr Lys Ser Lys
145 150 155 160

Ile Ser Gly Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys Gly
165 170 175

Lys Pro Val Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu Gly
180 185 190

Ser Pro Val Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu Phe
195 200 205

Lys Ser Phe Asp Asp Phe Arg Lys Lys Phe Trp Glu Glu Val Ser Lys
210 215 220

Asp Pro Glu Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg Met
225 230 235 240

Lys Val Gly Lys Ala Pro Lys Thr Arg Thr Gln Asp Val Ser Gly Lys
245 250 255

Arg Thr Ser Phe Glu Leu His His Glu Lys Pro Ile Ser Gln Asn Gly
260 265 270

Gly Val Tyr Asp Met Asp Asn Ile Ser Val Val Thr Pro Lys Arg His
275 280 285

Ile Asp Ile His Arg Gly Lys Gly Ser Ser Ala Asp
290 295 300

<210> 134

<211> 518
<212> PRT
<213> artificial sequence

<220>
<223> sT1

<400> 134

```
Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu Leu Pro
1               5                   10                  15

Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly Val Ser
            20                  25                  30

Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg Thr Met
        35                  40                  45

Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala Phe Ser
    50                  55                  60

Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser Leu Phe
65                  70                  75                  80

Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His His Thr
                85                  90                  95

Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu Arg Ala
            100                 105                 110

Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala Ala Arg
        115                 120                 125

Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln Pro Ser
    130                 135                 140

Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr Ser
145                 150                 155                 160

Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala
                165                 170                 175

Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile
            180                 185                 190

Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys
        195                 200                 205

Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile
    210                 215                 220

Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu
225                 230                 235                 240

Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr
                245                 250                 255

Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val Glu
            260                 265                 270
```

```
Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn Ser
        275                 280                 285

Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu Thr
    290                 295                 300

Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu
305                 310                 315                 320

Asp Ala Val Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile Lys Arg
                325                 330                 335

Thr Asn Arg Arg Ile Pro Glu Arg Thr Ser His Arg Val Ala Asp His
            340                 345                 350

Ala Gln Val Val Arg Val Leu Gly Phe Phe Gln Cys His Ser His Pro
        355                 360                 365

Ala Gln Ala Phe Asp Asp Ala Met Thr Gln Phe Gly Met Ser Arg His
    370                 375                 380

Gly Leu Leu Gln Leu Phe Arg Arg Val Gly Val Thr Glu Leu Glu Ala
385                 390                 395                 400

Arg Ser Gly Thr Leu Pro Pro Ala Ser Gln Arg Trp Asp Arg Ile Leu
                405                 410                 415

Gln Ala Ser Gly Met Lys Arg Ala Lys Pro Ser Pro Thr Ser Thr Gln
            420                 425                 430

Thr Pro Asp Gln Ala Ser Leu His Ala Phe Ala Asp Ser Leu Glu Arg
        435                 440                 445

Asp Leu Asp Ala Pro Ser Pro Met His Glu Gly Asp Gln Thr Arg Ala
    450                 455                 460

Ser Ser Arg Lys Arg Ser Arg Ser Asp Arg Ala Val Thr Gly Pro Ser
465                 470                 475                 480

Ala Gln Gln Ser Phe Glu Val Arg Val Pro Glu Gln Arg Asp Ala Leu
                485                 490                 495

His Leu Pro Leu Ser Trp Arg Val Lys Arg Pro Arg Thr Ser Ile Gly
            500                 505                 510

Gly Gly Leu Pro Asp Pro
        515
```

<210> 135

<211> 147

<212> PRT
<213> artificial sequence

<220>
<223> sT2

<400> 135
Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile
1 5 10 15

Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val
20 25 30

Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro
35 40 45

Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala
50 55 60

Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp
65 70 75 80

Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu
85 90 95

Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala
100 105 110

Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn
115 120 125

Ala Leu Thr Gly Ala Pro Leu Asn Ser Ile Val Ala Gln Leu Ser Arg
130 135 140

Pro Asp Pro
145

<210> 136

<211> 1116
<212> PRT
<213> artificial sequence

<220>
<223> bT1-Avr

<400> 136
Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu Leu Pro
1 5 10 15

Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly Val Ser
20 25 30

Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg Thr Met
35 40 45

```
Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala Phe Ser
    50                  55                  60

Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser Leu Phe
65                  70                  75                  80

Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His His Thr
                85                  90                  95

Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu Arg Ala
            100                 105                 110

Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala Ala Arg
        115                 120                 125

Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln Pro Ser
    130                 135                 140

Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr Ser
145                 150                 155                 160

Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala
                165                 170                 175

Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile
            180                 185                 190

Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys
        195                 200                 205

Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile
    210                 215                 220

Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu
225                 230                 235                 240

Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr
                245                 250                 255

Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val Glu
            260                 265                 270

Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu
        275                 280                 285

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
    290                 295                 300

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
305                 310                 315                 320
```

```
Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly
                325                 330                 335

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            340                 345                 350

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile
        355                 360                 365

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu
    370                 375                 380

Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser
385                 390                 395                 400

Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
                405                 410                 415

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
            420                 425                 430

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
        435                 440                 445

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
    450                 455                 460

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
465                 470                 475                 480

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                485                 490                 495

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
            500                 505                 510

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        515                 520                 525

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
    530                 535                 540

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
545                 550                 555                 560

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
                565                 570                 575

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            580                 585                 590
```

Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly
595 600 605

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
610 615 620

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile
625 630 635 640

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu
645 650 655

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
660 665 670

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
675 680 685

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
690 695 700

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
705 710 715 720

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
725 730 735

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
740 745 750

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
755 760 765

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
770 775 780

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
785 790 795 800

Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu
805 810 815

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
820 825 830

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
835 840 845

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
850 855 860

Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly
865 870 875 880

Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro
885 890 895

Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu
900 905 910

Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala
915 920 925

Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg Thr Ser
930 935 940

His Arg Val Ala Asp His Ala Gln Val Val Arg Val Leu Gly Phe Phe
945 950 955 960

Gln Cys His Ser His Pro Ala Gln Ala Phe Asp Asp Ala Met Thr Gln
965 970 975

Phe Gly Met Ser Arg His Gly Leu Leu Gln Leu Phe Arg Arg Val Gly
980 985 990

Val Thr Glu Leu Glu Ala Arg Ser Gly Thr Leu Pro Pro Ala Ser Gln
995 1000 1005

Arg Trp Asp Arg Ile Leu Gln Ala Ser Gly Met Lys Arg Ala Lys Pro
1010 1015 1020

Ser Pro Thr Ser Thr Gln Thr Pro Asp Gln Ala Ser Leu His Ala Phe
1025 1030 1035 1040

Ala Asp Ser Leu Glu Arg Asp Leu Asp Ala Pro Ser Pro Met His Glu
1045 1050 1055

Gly Asp Gln Thr Arg Ala Ser Ser Arg Lys Arg Ser Arg Ser Asp Arg
1060 1065 1070

Ala Val Thr Gly Pro Ser Ala Gln Gln Ser Phe Glu Val Arg Val Pro
1075 1080 1085

Glu Gln Arg Asp Ala Leu His Leu Pro Leu Ser Trp Arg Val Lys Arg
1090 1095 1100

Pro Arg Thr Ser Ile Gly Gly Gly Leu Pro Asp Pro
1105 1110 1115

<210> 137

<211> 745

<212> PRT
<213> artificial sequence

<220>
<223> bT2-Avr

<400> 137

```
Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile
1               5                  10                  15

Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val
            20                  25                  30

Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro
        35                  40                  45

Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala
    50                  55                  60

Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp
65                  70                  75                  80

Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu
                85                  90                  95

Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala
            100                 105                 110

Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn
        115                 120                 125

Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala
    130                 135                 140

Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala
145                 150                 155                 160

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
                165                 170                 175

Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val
            180                 185                 190

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
        195                 200                 205

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu
    210                 215                 220

Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
225                 230                 235                 240

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala
```

245 250 255

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
260 265 270

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys
275 280 285

Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala
290 295 300

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly
305 310 315 320

Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys
325 330 335

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn
340 345 350

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val
355 360 365

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
370 375 380

Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
385 390 395 400

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
405 410 415

Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
420 425 430

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
435 440 445

Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val
450 455 460

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
465 470 475 480

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu
485 490 495

Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
500 505 510

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala

```
        515                 520                 525

Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly
    530                 535                 540

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
545                 550                 555                 560

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
                565                 570                 575

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly
            580                 585                 590

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
        595                 600                 605

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His
    610                 615                 620

Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
625                 630                 635                 640

Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
                645                 650                 655

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
            660                 665                 670

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
        675                 680                 685

Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
    690                 695                 700

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
705                 710                 715                 720

Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile
                725                 730                 735

Val Ala Gln Leu Ser Arg Pro Asp Pro
            740                 745
```

<210> 138

<211> 1320
<212> PRT
<213> artificial sequence

<220>
<223> bT1-Pth

<400> 138

```
Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu Leu Pro
1               5                   10                  15

Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly Val Ser
            20                  25                  30

Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg Thr Met
        35                  40                  45

Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala Phe Ser
    50                  55                  60

Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser Leu Phe
65                  70                  75                  80

Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His His Thr
                85                  90                  95

Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu Arg Ala
            100                 105                 110

Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala Ala Arg
        115                 120                 125

Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln Pro Ser
    130                 135                 140

Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr Ser
145                 150                 155                 160

Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala
                165                 170                 175

Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile
            180                 185                 190

Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys
        195                 200                 205

Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile
    210                 215                 220

Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu
225                 230                 235                 240

Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr
                245                 250                 255

Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val Glu
            260                 265                 270
```

```
Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu
        275                 280                 285

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Lys Gly Gly Lys Gln
    290                 295                 300

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
305                 310                 315                 320

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
                325                 330                 335

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            340                 345                 350

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile
        355                 360                 365

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu
    370                 375                 380

Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser
385                 390                 395                 400

Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
                405                 410                 415

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
            420                 425                 430

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
        435                 440                 445

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val
    450                 455                 460

Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
465                 470                 475                 480

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                485                 490                 495

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
            500                 505                 510

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        515                 520                 525

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
    530                 535                 540
```

```
Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
545                 550                 555                 560

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
                565                 570                 575

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            580                 585                 590

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
        595                 600                 605

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    610                 615                 620

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
625                 630                 635                 640

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
                645                 650                 655

Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser
            660                 665                 670

Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
        675                 680                 685

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
    690                 695                 700

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
705                 710                 715                 720

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
                725                 730                 735

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
            740                 745                 750

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
        755                 760                 765

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
    770                 775                 780

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
785                 790                 795                 800

Glu Gln Val Val Ala Ile Ala Ser Asn Lys Gly Gly Lys Gln Ala Leu
                805                 810                 815
```

```
Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
            820                 825                 830

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
        835                 840                 845

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
    850                 855                 860

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
865                 870                 875                 880

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
                885                 890                 895

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
            900                 905                 910

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
        915                 920                 925

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
    930                 935                 940

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
945                 950                 955                 960

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
                965                 970                 975

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
            980                 985                 990

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
        995                 1000                1005

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
    1010                1015                1020

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
1025                1030                1035                1040

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
                1045                1050                1055

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
            1060                1065                1070

Gln Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Arg Pro Ala Leu
        1075                1080                1085
```

Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala
1090 1095 1100

Leu Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro
1105 1110 1115 1120

Ala Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile
1125 1130 1135

Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg Thr Ser His Arg Val Ala
1140 1145 1150

Asp His Ala Gln Val Val Arg Val Leu Gly Phe Phe Gln Cys His Ser
1155 1160 1165

His Pro Ala Gln Ala Phe Asp Asp Ala Met Thr Gln Phe Gly Met Ser
1170 1175 1180

Arg His Gly Leu Leu Gln Leu Phe Arg Arg Val Gly Val Thr Glu Leu
1185 1190 1195 1200

Glu Ala Arg Ser Gly Thr Leu Pro Pro Ala Ser Gln Arg Trp Asp Arg
1205 1210 1215

Ile Leu Gln Ala Ser Gly Met Lys Arg Ala Lys Pro Ser Pro Thr Ser
1220 1225 1230

Thr Gln Thr Pro Asp Gln Ala Ser Leu His Ala Phe Ala Asp Ser Leu
1235 1240 1245

Glu Arg Asp Leu Asp Ala Pro Ser Pro Met His Glu Gly Asp Gln Thr
1250 1255 1260

Arg Ala Ser Ser Arg Lys Arg Ser Arg Ser Asp Arg Ala Val Thr Gly
1265 1270 1275 1280

Pro Ser Ala Gln Gln Ser Phe Glu Val Arg Val Pro Glu Gln Arg Asp
1285 1290 1295

Ala Leu His Leu Pro Leu Ser Trp Arg Val Lys Arg Pro Arg Thr Ser
1300 1305 1310

Ile Gly Gly Gly Leu Pro Asp Pro
1315 1320

<210> 139

<211> 949
<212> PRT
<213> artificial sequence

<220>

<223> bT2-Pth

<400> 139

```
Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile
1               5                  10                  15

Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val
            20                  25                  30

Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro
        35                  40                  45

Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala
    50                  55                  60

Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp
65                  70                  75                  80

Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu
                85                  90                  95

Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala
            100                 105                 110

Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn
        115                 120                 125

Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala
    130                 135                 140

Ile Ala Ser Asn Lys Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
145                 150                 155                 160

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
                165                 170                 175

Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val
            180                 185                 190

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
        195                 200                 205

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu
    210                 215                 220

Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
225                 230                 235                 240

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala
                245                 250                 255

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
```

260 265 270

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
275 280 285

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
290 295 300

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
305 310 315 320

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
325 330 335

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His
340 345 350

Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
355 360 365

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
370 375 380

Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
385 390 395 400

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
405 410 415

Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
420 425 430

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
435 440 445

Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val
450 455 460

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
465 470 475 480

Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu
485 490 495

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
500 505 510

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala
515 520 525

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly

530 535 540

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys
545 550 555 560

Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala
565 570 575

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly
580 585 590

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
595 600 605

Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
610 615 620

Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
625 630 635 640

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
645 650 655

Ser Asn Lys Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
660 665 670

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
675 680 685

Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
690 695 700

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
705 710 715 720

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val
725 730 735

Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
740 745 750

Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu
755 760 765

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
770 775 780

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala
785 790 795 800

Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly

805 810 815

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
820 825 830

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
835 840 845

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly
850 855 860

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
865 870 875 880

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn
885 890 895

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val
900 905 910

Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
915 920 925

Ser His Asp Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu
930 935 940

Ser Arg Pro Asp Pro
945

<210> 140

<211> 141
<212> PRT
<213> artificial sequence

<220>
<223> ColE7

<400> 140
Met Ala Ser Gly Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys
1 5 10 15

Gly Lys Pro Val Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu
20 25 30

Gly Ser Pro Val Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu
35 40 45

Phe Lys Ser Phe Asp Asp Phe Arg Lys Lys Phe Trp Glu Glu Val Ser
50 55 60

Lys Asp Pro Glu Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg
65 70 75 80

```
Met Lys Val Gly Lys Ala Pro Lys Thr Arg Thr Gln Asp Val Ser Gly
                85                  90                  95

Lys Arg Thr Ser Phe Glu Leu His His Glu Lys Pro Ile Ser Gln Asn
            100                 105                 110

Gly Gly Val Tyr Asp Met Asp Asn Ile Ser Val Val Thr Pro Lys Arg
        115                 120                 125

His Ile Asp Ile His Arg Gly Lys Gly Ser Ser Ala Asp
    130                 135                 140
```

<210> 141

<211> 2217
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS7183 Scaffold

<400> 141

```
atggccgacc ccattcgttc gcgcacacca agtcctgccc gcgagcttct gcccggaccc     60
caacccgatg gggttcagcc gactgcagat cgtggggtgt ctccgcctgc cggcggcccc    120
ctggatggct tgccggctcg gcggacgatg tcccggaccc ggctgccatc tccccctgcc    180
ccctcacctg cgttctcggc gggcagcttc agtgacctgt tacgtcagtt cgatccgtca    240
ctttttaata catcgctttt tgattcattg cctcccttcg gcgctcacca tacagaggct    300
gccacaggcg agtgggatga ggtgcaatcg ggtctgcggg cagccgacgc ccccccaccc    360
accatgcgcg tggctgtcac tgccgcgcgg cccccgcgcg ccaagccggc gccgcgacga    420
cgtgctgcgc aaccctccga cgcttcgccg gcggcgcagg tggatctacg cacgctcggc    480
tacagccagc agcaacagga gaagatcaaa ccgaaggttc gttcgacagt ggcgcagcac    540
cacgaggcac tggtcggcca cgggtttaca cacgcgcaca tcgttgcgtt aagccaacac    600
ccggcagcgt tagggaccgt cgctgtcaag tatcaggaca tgatcgcagc gttgccagag    660
gcgacacacg aagcgatcgt tggcgtcggc aaacagtggt ccggcgcacg cgctctggag    720
gccttgctca cggtggcggg agagttgaga ggtccaccgt tacagttgga cacaggccaa    780
cttctcaaga ttgcaaaacg tggcggcgtg accgcagtgg aggcagtgca tgcatggcgc    840
aatgcactga cgggtgcccc gctcaacttg accggagacg cccgggggat caggtcacgt    900
gcgtctcgga gcattgttgc ccagttatct cgccctgatc cggcgttggc cgcgttgacc    960
aacgaccacc tcgtcgcctt ggcctgcctc ggcgggcgtc ctgcgctgga tgcagtgaaa   1020
aagggattgc cgcacgcgcc ggccttgatc aaaagaacca atcgccgtat tcccgaacgc   1080
acatcccatc gcgttgccga ccacgcgcaa gtggttcgcg tgctgggttt tttccagtgc   1140
cactcccacc cagcgcaagc atttgatgac gccatgacgc agttcgggat gagcaggcac   1200
gggttgttac agctctttcg cagagtgggc gtcaccgaac tcgaagcccg cagtggaacg   1260
ctccccccag ccagtcagcg ttgggaccgt atcctccagg catcagggat gaaaagggcc   1320
aaaccgtccc ctacttcaac tcaaacgccg gatcaggcgt ctttgcatgc attcgccgat   1380
tcgctggagc gtgaccttga tgcgcctagc ccaatgcacg agggagatca gacgcgggca   1440
agtagccgta aacggtcccg atcggatcgt gctgtcaccg gtccctccgc acagcaatcg   1500
ttcgaggtgc gcgttcccga acagcgcgat gcgctccatt tgcccctctc ctggagggta   1560
aaacgcccgc gtaccagtat cgggggcggc ctcccggatc ctatcagccg ttcccagctg   1620
gtgaagtccg agctggagga gaagaaatcc gagttgaggc acaagctgaa gtacgtgccc   1680
cacgagtaca tcgagctgat cgagatcgcc cggaacagca cccaggaccg tatcctggag   1740
atgaaggtga tggagttctt catgaaggtg tacggctaca ggggcaagca cctgggcggc   1800
tccaggaagc ccgacggcgc catctacacc gtgggctccc ccatcgacta cggcgtgatc   1860
gtggacacca aggcctactc cggcggctac aacctgccca tcggccaggc cgacgaaatg   1920
cagaggtacg tggaggagaa ccagaccagg aacaagcaca tcaaccccaa cgagtggtgg   1980
aaggtgtacc cctccagcgt gaccgagttc aagttcctgt tcgtgtccgg ccacttcaag   2040
ggcaactaca aggcccagct gaccaggctg aaccacatca ccaactgcaa cggcgccgtg   2100
ctgtccgtgg aggagctcct gatcggcggc gagatgatca aggccggcac cctgaccctg   2160
gaggaggtga ggaggaagtt caacaacggc gagatcaact tcgcggccga ctgataa      2217
```

<210> 142

<211> 35
<212> DNA
<213> Artificial sequence

<220>
<223> Primer CMP_G061

<400> 142
cggcggccat ggtggatcta cgcacgctcg gctac 35

<210> 143

<211> 36
<212> DNA
<213> Artificial sequence

<220>
<223> Primer CMP_G065

<400> 143
cgcggcggat cccggatcag ggcgagataa ctgggc 36

<210> 144

<211> 3032
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS7865

<400> 144
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 60
tggccttttg ctcacatggt ctttcctgcg ttatcccctg attctgtgga taaccgtatt 120
accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca 180
gtgagcgagg aagcggagag cgcccaatac gcaaaccgcc tctccccgcg cgttggccga 240
ttcattaatg cagctggcac gacaggtttc ccgactggaa agcgggcagt gagcgcaacg 300
caattaatgt gagttagctc actcattagg caccccaggc tttacacttt atgcttccgg 360
ctcgtatgtt gtgtggaatt gtgagcggat aacaatttca cacaggaaac agctatgacc 420
atgattacgc caagcgcgtc aattaaccct cactaaaggg aacaaaagct gttaattaag 480
cgcgccacac gcaaacacaa atacacagcg gccttgccac catggcaaac accggtggat 540
cctccggaag atcttcggcc gactgataag tcgaccacca ccaccaccac cactgataag 600
agctcctcga gcctgcagca gctgaagctt tggacttctg gcgcgcccaa ttcgccctat 660
agtgagtcgt attacgtcgc gctcactggc cgtcgtttta caacgtcgtg actgggaaaa 720
ccctggcgtt acccaactta atcgccttgc agcacatccc cctttcgcca gctggcgtaa 780
tagcgaagag gcccgcaccg aaacgccctt cccaacagtt gcgcagcctg aatggcgaat 840
gggagcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac 900
cgctacactt gccagcgccc tagcgcccgc tcctttcgct ttcttccctt cctttctcgc 960
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt 1020
tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatggtt ggcctgtagt 1080
gggccatagc cctgatagac ggtttttcgc cctttgacgt tggagtccac gttctttaat 1140
agtggactct tgttccaaac tggaacaaca ctcaacccta tctcggtcta ttcttttgat 1200
ttataaggga ttttgccgat ttcggcctat tggttaaaaa atgagctgat ttaacaaaaa 1260
tttaacgcga attttaacaa aatattaacg cttacaattt aggtggcact tttcggggaa 1320
atgtgcgcgg aacccctatt tgtttatttt tctaaataca ttcaaatatg tatccgctca 1380
tgagacaata accctgataa atgcttcaat aatattgaaa aaggaagagt atgagtattc 1440
aacatttccg tgtcgccctt attccctttt ttgcggcatt ttgccttcct gtttttgctc 1500
acccagaaac gctggtgaaa gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt 1560
acatcgaact ggatctcaac agcggtaaga tccttgagag ttttcgcccc gaagaacgtt 1620
ttccaatgat gagcactttt aaagttctgc tatgtggcgc ggtattatcc cgtattgacg 1680
ccgggcaaga gcaactcggt cgccgcatac actattctca gaatgacttg gttgagtact 1740
caccagtcac agaaaagcat cttacggatg gcatgacagt aagagaatta tgcagtgctg 1800
ccataaccat gagtgataac actgcggcca acttacttct gacaacgatc ggaggaccga 1860
aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgcctt gatcgttggg 1920
aaccggagct gaatgaagcc ataccaaacg acgagcgtga caccacgatg cctgtagcaa 1980
tggcaacaac gttgcgcaaa ctattaactg gcgaactact tactctagct tcccggcaac 2040

aattaataga ctggatggag gcggataaag ttgcaggacc acttctgcgc tcggcccttc 2100
cggctggctg gtttattgct gataaatctg gagccggtga gcgtggttct cgcggtatca 2160
ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac acgacgggga 2220
gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc tcactgatta 2280
agcattggta actgtcagac caagtttact catatatact ttagattgat ttaaaacttc 2340
atttttaatt taaaaggatc taggtgaaga tccttttttga taatctcatg accaaaatcc 2400
cttaacgtga gttttcgttc cactgagcgt cagaccccgt agaaaagatc aaaggatctt 2460
cttgagatcc ttttttttctg cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac 2520
cagcggtggt ttgtttgccg gatcaagagc taccaactct ttttccgaag gtaactggct 2580
tcagcagagc gcagatacca aatactgttc ttctagtgta gccgtagtta ggccaccact 2640
tcaagaactc tgtagcaccg cctacatacc tcgctctgct aatcctgtta ccagtggctg 2700
ctgccagtgg cgataagtcg tgtcttaccg ggttggactc aagacgatag ttaccggata 2760
aggcgcagcg gtcgggctga acgggggggtt cgtgcacaca gcccagcttg gagcgaacga 2820
cctacaccga actgagatac ctacagcgtg agctatgaga aagcgccacg cttcccgaag 2880
ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg 2940
agcttccagg gggaaacgcc tggtatcttt atagtcctgt cgggtttcgc cacctctgac 3000
ttgagcgtcg atttttgtga tgctcgtcag gg 3032

<210> 145

<211> 3503
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS9009 (pCLS7865-cTAL11_CFS1)

<400> 145
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 60
tggccttttg ctcacatggt ctttcctgcg ttatcccctg attctgtgga taaccgtatt 120
accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca 180
gtgagcgagg aagcggagag cgcccaatac gcaaaccgcc tctccccgcg cgttggccga 240
ttcattaatg cagctggcac gacaggtttc ccgactggaa agcgggcagt gagcgcaacg 300
caattaatgt gagttagctc actcattagg caccccaggc tttacacttt atgcttccgg 360
ctcgtatgtt gtgtggaatt gtgagcggat aacaatttca cacaggaaac agctatgacc 420
atgattacgc caagcgcgtc aattaaccct cactaaaggg aacaaaagct gttaattaag 480
cgcgccacac gcaaacacaa atacacagcg gccttgccac catggtggat ctacgcacgc 540
tcggctacag ccagcagcaa caggagaaga tcaaaccgaa ggttcgttcg acagtggcgc 600
agcaccacga ggcactggtc ggccacgggt ttacacacgc gcacatcgtt gcgttaagcc 660
aacacccggc agcgttaggg accgtcgctg tcaagtatca ggacatgatc gcagcgttgc 720
cagaggcgac acacgaagcg atcgttggcg tcggcaaaca gtggtccggc gcacgcgctc 780
tggaggcctt gctcacggtg gcgggagagt tgagaggtcc accgttacag ttggacacag 840
gccaacttct caagattgca aaacgtggcg gcgtgaccgc agtggaggca gtgcatgcat 900
ggcgcaatgc actgacgggt gccccgctca acttgaccgg agacgcccgg gggatcaggt 960
cacgtgcgtc tcggagcatt gttgcccagt tatctcgccc tgatccggga tcctccggaa 1020
gatcttcggc cgactgataa gtcgaccacc accaccacca ccactgataa gagctcctcg 1080
agcctgcagc agctgaagct ttggacttct ggcgcgccca attcgcccta tagtgagtcg 1140
tattacgtcg cgctcactgg ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt 1200
tacccaactt aatcgccttg cagcacatcc ccctttcgcc agctggcgta atagcgaaga 1260
ggcccgcacc gaaacgccct tcccaacagt tgcgcagcct gaatggcgaa tgggagcgcc 1320
ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg cgcagcgtga ccgctacact 1380
tgccagcgcc ctagcgcccg ctcctttcgc tttcttccct tcctttctcg ccacgttcgc 1440
cggctttccc cgtcaagctc taaatcgggg gctcccttta gggttccgat ttagtgcttt 1500
acggcacctc gaccccaaaa aacttgatta gggtgatggt tggcctgtag tgggccatag 1560
ccctgataga cggtttttcg ccctttgacg ttggagtcca cgttctttaa tagtggactc 1620
ttgttccaaa ctggaacaac actcaaccct atctcggtct attcttttga tttataaggg 1680
attttgccga tttcggccta ttggttaaaa aatgagctga tttaacaaaa atttaacgcg 1740
aattttaaca aaatattaac gcttacaatt taggtggcac ttttcgggga aatgtgcgcg 1800
gaacccctat ttgtttattt ttctaaatac attcaaatat gtatccgctc atgagacaat 1860
aaccctgata aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacatttcc 1920
gtgtcgccct tattcccttt tttgcggcat tttgccttcc tgtttttgct cacccagaaa 1980
cgctggtgaa agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgaac 2040
tggatctcaa cagcggtaag atccttgaga gttttcgccc cgaagaacgt tttccaatga 2100
tgagcacttt taaagttctg ctatgtggcg cggtattatc ccgtattgac gccgggcaag 2160
agcaactcgg tcgccgcata cactattctc agaatgactt ggttgagtac tcaccagtca 2220
cagaaaagca tcttacggat ggcatgacag taagagaatt atgcagtgct gccataacca 2280
tgagtgataa cactgcggcc aacttacttc tgacaacgat cggaggaccg aaggagctaa 2340

ccgctttttt gcacaacatg ggggatcatg taactcgcct tgatcgttgg gaaccggagc 2400
tgaatgaagc cataccaaac gacgagcgtg acaccacgat gcctgtagca atggcaacaa 2460
cgttgcgcaa actattaact ggcgaactac ttactctagc ttcccggcaa caattaatag 2520
actggatgga ggcggataaa gttgcaggac cacttctgcg ctcggccctt ccggctggct 2580
ggtttattgc tgataaatct ggagccggtg agcgtggttc tcgcggtatc attgcagcac 2640
tggggccaga tggtaagccc tcccgtatcg tagttatcta cacgacgggg agtcaggcaa 2700
ctatggatga acgaaataga cagatcgctg agataggtgc ctcactgatt aagcattggt 2760
aactgtcaga ccaagtttac tcatatatac tttagattga tttaaaactt catttttaat 2820
ttaaaaggat ctaggtgaag atccttttg ataatctcat gaccaaaatc ccttaacgtg 2880
agttttcgtt ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc 2940
ctttttttct gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcggtgg 3000
tttgtttgcc ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag 3060
cgcagatacc aaatactgtt cttctagtgt agccgtagtt aggccaccac ttcaagaact 3120
ctgtagcacc gcctacatac ctcgctctgc taatcctgtt accagtggct gctgccagtg 3180
gcgataagtc gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc 3240
ggtcgggctg aacgggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg 3300
aactgagata cctacagcgt gagctatgag aaagcgccac gcttcccgaa gggagaaagg 3360
cggacaggta tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag 3420
ggggaaacgc ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc 3480
gatttttgtg atgctcgtca ggg 3503

<210> 146

<211> 9402
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS6614

<400> 146
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 60
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 120
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 180
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 240
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 300
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 360
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 420
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 480
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 540
tttttgtttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 600
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 660
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 720
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 780
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 840
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 900
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 960
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1020
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1080
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1140
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1200
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1260
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1320
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1380
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1440
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1500
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1560
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1620
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1680
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1740
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1800
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1860
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1920
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 1980
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2040
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2100
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2160

```
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca   2220
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc   2280
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata   2340
gacacgcaaa cacaaataca cagcggcctt gccaccatgg ccaagtctgg catctaccag   2400
attaagaaca ccctgaataa caaagtgtat gttggatctg ccaaggactt tgagaaaaga   2460
tggaagaggc acttcaagga tctggaaaag ggttgccaca gctccatcaa gctccaaagg   2520
tctttcaaca aacatggcaa tgtctttgag tgctccattt tggaagagat cccttatgaa   2580
aaggacctca ttatagagag ggaaaacttc tggatcaagg agctgaactc aaaaatcaat   2640
ggatacaaca ttgcagatgc cacatttggt gacacctgtt ctactcaccc ccttaaggaa   2700
gagatcatta aaaagaggtc tgagactgtg aaggctaaga tgttgaaact ggggcctgat   2760
ggcagaaagg cactctactc taagccaggt agcaaaaatg gaaggtggaa ccctgagacc   2820
cacaagttct gcaaatgtgg ggtgagaata cagacatcag cctacacttg cagcaagtgt   2880
aggaacagat caggagagaa taactccttt ttcaaccaca agcattctga catcaccaag   2940
tccaaaattt caggcaagat gaaaggtcaa ggtccatctg gcaataccaa atataacaaa   3000
gagttcctgc tgtacctggc cggctttgtg gacggtaacg gtagcatcat cgctcagatt   3060
aaaccaaacc agtcttataa gtttaaacat cagctaagct tgacctttca ggtgactcaa   3120
aagacccagc gccgttggtt tctggacaaa ctagtggatg aaattggcgt tggttacgta   3180
cgtgatcgcg gatccgtttc cgattacatt ttaagcgaaa tcaagccgct gcacaacttc   3240
ctgactcaac tgcagccgtt tctgaaactg aaacagaaac aggcaaacct ggttctgaaa   3300
attatcgaac agctgccgtc tgcaaaagaa tccccggaca aattcctgga agtttgtacc   3360
tgggtggatc agattgcagc tctgaacgat tctaagacgc gtaaaaccac ttctgaaacc   3420
gttcgtgctg tgctggacag cctgagcgag aagaagaaat cctccccggc ggccgactga   3480
taactcgagc gatcctctag acgagctcct cgagcctgca gcagctgaag ctttggactt   3540
cttcgccaga ggtttggtca agtctccaat caaggttgtc ggcttgtcta ccttgccaga   3600
aatttacgaa aagatggaaa agggtcaaat cgttggtaga tacgttgttg acacttctaa   3660
ataagcgaat ttcttatgat ttatgatttt tattattaaa taagttataa aaaaaataag   3720
tgtatacaaa ttttaaagtg actcttaggt tttaaaacga aaattcttat tcttgagtaa   3780
ctctttcctg taggtcaggt tgctttctca ggtatagcat gaggtcgctc ttattgacca   3840
cacctctacc ggcatgcaag cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat   3900
tgttatccgc tcacaattcc acacaacata cgagccggaa gcataaagtg taaagcctgg   3960
ggtgcctaat gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag   4020
tcgggaaacc tgtcgtgcca gcagatctat tacattatgg gtggtatgtt ggaataaaaa   4080
tcaactatca tctactaact agtatttacg ttactagtat attatcatat acggtgttag   4140
aagatgacgc aaatgatgag aaatagtcat ctaaattagt ggaagctgaa acgcaaggat   4200
tgataatgta ataggatcaa tgaatattaa catataaaat gatgataata atatttatag   4260
aattgtgtag aattgcagat tccctttttat ggattcctaa atcctcgagg agaacttcta   4320
gtatatctac atacctaata ttattgcctt attaaaaatg gaatcccaac aattacatca   4380
aaatccacat tctcttcaaa atcaattgtc ctgtacttcc ttgttcatgt gtgttcaaaa   4440
acgttatatt tataggataa ttatactcta tttctcaaca agtaattggt tgtttggccg   4500
agcggtctaa ggcgcctgat tcaagaaata tcttgaccgc agttaactgt gggaatactc   4560
aggtatcgta agatgcaaga gttcgaatct cttagcaacc attatttttt tcctcaacat   4620
aacgagaaca cacaggggcg ctatcgcaca gaatcaaatt cgatgactgg aaatttttg   4680
ttaatttcag aggtcgcctg acgcatatac ctttttcaac tgaaaaattg ggagaaaaag   4740
gaaaggtgag agccgcggaa ccggcttttc atatagaata gagaagcgtt catgactaaa   4800
tgcttgcatc acaatacttg aagttgacaa tattatttaa ggacctattg ttttttccaa   4860
taggtggtta gcaatcgtct tactttctaa cttttcttac cttttacatt tcagcaatat   4920
atatatatat atttcaagga tataccattc taatgtctgc ccctaagaag atcgtcgttt   4980
tgccaggtga ccacgttggt caagaaatca cagccgaagc cattaaggtt cttaaagcta   5040
tttctgatgt tcgttccaat gtcaagttcg atttcgaaaa tcatttaatt ggtggtgctg   5100
ctatcgatgc tacaggtgtc ccacttccag atgaggcgct ggaagcctcc aagaaggttg   5160
atgccgtttt gttaggtgct gtgggtggtc ctaaatgggg taccggtagt gttagacctg   5220
aacaaggttt actaaaaatc cgtaaagaac ttcaattgta cgccaactta agaccatgta   5280
actttgcatc cgactctctt ttagacttat ctccaatcaa gccacaattt gctaaaggta   5340
ctgacttcgt tgttgtcaga gaattagtgg gaggtattta ctttggtaag agaaaggaag   5400
acgatggtga tggtgtcgct tgggatagtg aacaatacac cgttccagaa gtgcaaagaa   5460
tcacaagaat ggccgctttc atggccctac aacatgagcc accattgcct atttggtcct   5520
tggataaagc taatgttttg gcctcttcaa gattatggag aaaaactgtg gaggaaacca   5580
tcaagaacga attccctaca ttgaaggttc aacatcaatt gattgattct gccgccatga   5640
tcctagttaa gaacccaacc cacctaaatg gtattataat caccagcaac atgtttggtg   5700
atatcatctc cgatgaagcc tccgttatcc caggttcctt gggtttgttg ccatctgcgt   5760
ccttggcctc tttgccagac aagaacaccg catttggttt gtacgaacca tgccacggtt   5820
ctgctccaga tttgccaaag aataaggtca accctatcgc cactatcttg tctgctgcaa   5880
tgatgttgaa attgtcattg aacttgcctg aagaaggtaa ggccattgaa gatgcagtta   5940
aaaaggtttt ggatgcaggt atcagaactg gtgatttagg tggttccaac agtaccacgg   6000
aagtcggtga tgctgtcgcc gaagaagtta agaaaatcct tgcttaaaaa gattctcttt   6060
ttttatgata tttgtacata aactttataa atgaaattca taatagaaac gacacgaaat   6120
tacaaaatgg aatatgttca tagggtagac gaaactatat acgcaatcta catacattta   6180
tcaagaagga gaaaaaggag gatgtaaagg aatacaggta agcaaattga tactaatggc   6240
```

tcaacgtgat aaggaaaaag aattgcactt taacattaat attgacaagg aggagggcac 6300
cacacaaaaa gttaggtgta acagaaaatc atgaaactat gattcctaat ttatatattg 6360
gaggattttc tctaaaaaaa aaaaaataca acaaataaaa aacactcaat gacctgacca 6420
tttgatggag tttaagtcaa taccttcttg aaccatttcc cataatggtg aaagttccct 6480
caagaatttt actctgtcag aaacggcctt aacgacgtag tcgacctcct cttcagtact 6540
aaatctacca ataccaaatc tgatggaaga atgggctaat gcatcatcct tacccagcgc 6600
atgtaaaaca taagaaggtt ctagggaagc agatgtacag gctgaacccg aggataatgc 6660
gatatccctt agtgccatca ataaagattc tccttccacg taggcgaaag aaacgttaac 6720
acaccctgga taacgatgat ctggagatcc gttcaacgtg gtatgttcag cggataatag 6780
acctttgact aatttatcgg atagtctttt gatgtgagct tggtcgttgt caaattcttt 6840
cttcatcaat ctcgcagctt caccaaatcc cgctaccaat ggggggggca aagtaccaga 6900
tctgctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc 6960
ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc 7020
agcatcgatg aattccacgg actatagact atactagtat actccgtcta ctgtacgata 7080
cacttccgct caggtccttg tcctttaacg aggccttacc actcttttgt tactctattg 7140
atccagctca gcaaaggcag tgtgatctaa gattctatct tcgcgatgta gtaaaactag 7200
ctagaccgag aaagagacta gaaatgcaaa aggcacttct acaatggctg ccatcattat 7260
tatccgatgt gacgctgcag cttctcaatg atattcgaat acgctttgag gagatacagc 7320
ctaatatccg acaaactgtt ttacagattt acgatcgtac ttgttaccca tcattgaatt 7380
ttgaacatcc gaacctggga gttttccctg aaacagatag tatatttgaa cctgtataat 7440
aatatatagt ctagcgcttt acggaagaca atgtatgtat ttcggttcct ggagaaacta 7500
ttgcatctat tgcataggta atcttgcacg tcgcatcccc ggttcatttt ctgcgtttcc 7560
atcttgcact tcaatagcat atctttgtta acgaagcatc tgtgcttcat tttgtagaac 7620
aaaaatgcaa cgcgagagcg ctaatttttc aaacaaagaa tctgagctgc atttttacag 7680
aacagaaatg caacgcgaaa gcgctatttt accaacgaag aatctgtgct tcatttttgt 7740
aaaacaaaaa tgcaacgcga gagcgctaat ttttcaaaca aagaatctga gctgcatttt 7800
tacagaacag aaatgcaacg cgagagcgct attttaccaa caaagaatct atacttcttt 7860
tttgttctac aaaaatgcat cccgagagcg ctatttttct aacaaagcat cttagattac 7920
ttttttttctc ctttgtgcgc tctataatgc agtctcttga taacttttg cactgtaggt 7980
ccgttaaggt tagaagaagg ctactttggt gtctattttc tcttccataa aaaaagcctg 8040
actccacttc ccgcgtttac tgattactag cgaagctgcg ggtgcatttt ttcaagataa 8100
aggcatcccc gattatattc tataccgatg tggattgcgc atactttgtg aacagaaagt 8160
gatagcgttg atgattcttc attggtcaga aaattatgaa cggtttcttc tattttgtct 8220
ctatatacta cgtataggaa atgtttacat tttcgtattg ttttcgattc actctatgaa 8280
tagttcttac tacaattttt ttgtctaaag agtaatacta gagataaaca taaaaaatgt 8340
agaggtcgag tttagatgca agttcaagga gcgaaaggtg gatgggtagg ttatataggg 8400
atatagcaca gagatatata gcaaagagat actttgagc aatgtttgtg gaagcggtat 8460
tcgcaatatt ttagtagctc gttacagtcc ggtgcgtttt tggtttttg aaagtgcgtc 8520
ttcagagcgc ttttggtttt caaaagcgct ctgaagttcc tatactttct agagaatagg 8580
aacttcggaa taggaacttc aaagcgtttc cgaaaacgag cgcttccgaa aatgcaacgc 8640
gagctgcgca catacagctc actgttcacg tcgcacctat atctgcgtgt tgcctgtata 8700
tatatataca tgagaagaac ggcatagtgc gtgtttatgc ttaaatgcgt acttatatgc 8760
gtctatttat gtaggatgaa aggtagtcta gtacctcctg tgatattatc ccattccatg 8820
cggggtatcg tatgcttcct tcagcactac cctttagctg ttctatatgc tgccactcct 8880
caattggatt agtctcatcc ttcaatgcta tcatttcctt tgatattgga tcatatgcat 8940
agtaccgaga aactagtgcg aagtagtgat caggtattgc tgttatctga tgagtatacg 9000
ttgtcctggc cacggcagaa gcacgcttat cgctccaatt tcccacaaca ttagtcaact 9060
ccgttaggcc cttcattgaa agaaatgagg tcatcaaatg tcttccaatg tgagattttg 9120
ggccattttt tatagcaaag attgaataag gcgcattttt cttcaaagct ttattgtacg 9180
atctgactaa gttatctttt aataattggt attcctgttt attgcttgaa gaattgccgg 9240
tcctatttac tcgttttagg actggttcag aattcatcga tgctcactca aaggtcggta 9300
atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag 9360
caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt tt 9402

<210> 147

<211> 36
<212> DNA
<213> Artificial sequence

<220>
<223> Primer CMP_G069

<400> 147
gccaccggat ccaagtctgg catctaccag attaag 36

<210> 148

<211> 35
<212> DNA
<213> Artificial sequence

<220>
<223> Primer CMP_G070

<400> 148
tggtatcggc cgatggacct tgacctttca tcttg 35

<210> 149

<211> 35
<212> DNA
<213> Artificial sequence

<220>
<223> Primer CMP_G071

<400> 149
tggtatcggc cgatggacct tgagaatgct tgtgg 35

<210> 150

<211> 4085
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS9010 (pCLS 7865-cT11 TevD01)

<400> 150
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 60
tggccttttg ctcacatggt ctttcctgcg ttatcccctg attctgtgga taaccgtatt 120
accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca 180
gtgagcgagg aagcggagag cgcccaatac gcaaaccgcc tctccccgcg cgttggccga 240
ttcattaatg cagctggcac gacaggtttc ccgactggaa agcgggcagt gagcgcaacg 300
caattaatgt gagttagctc actcattagg caccccaggc tttacacttt atgcttccgg 360
ctcgtatgtt gtgtggaatt gtgagcggat aacaatttca cacaggaaac agctatgacc 420
atgattacgc caagcgcgtc aattaaccct cactaaaggg aacaaaagct gttaattaag 480
cgcgccacac gcaaacacaa atacacagcg gccttgccac catggtggat ctacgcacgc 540
tcggctacag ccagcagcaa caggagaaga tcaaaccgaa ggttcgttcg acagtggcgc 600
agcaccacga ggcactggtc ggccacgggt ttacacacgc gcacatcgtt gcgttaagcc 660
aacacccggc agcgttaggg accgtcgctg tcaagtatca ggacatgatc gcagcgttgc 720
cagaggcgac acacgaagcg atcgttggcg tcggcaaaca gtggtccggc gcacgcgctc 780
tggaggcctt gctcacggtg gcgggagagt tgagaggtcc accgttacag ttggacacag 840
gccaacttct caagattgca aaacgtggcg gcgtgaccgc agtggaggca gtgcatgcat 900
ggcgcaatgc actgacgggt gccccgctca acttgaccgg agacgcccgg gggatcaggt 960
cacgtgcgtc tcggagcatt gttgcccagt tatctcgccc tgatccggga tccaagtctg 1020
gcatctacca gattaagaac accctgaata acaaagtgta tgttggatct gccaaggact 1080
ttgagaaaag atggaagagg cacttcaagg atctggaaaa gggttgccac agctccatca 1140
agctccaaag gtctttcaac aaacatggca atgtctttga gtgctccatt ttggaagaga 1200
tcccttatga aaaggacctc attatagaga gggaaaactt ctggatcaag gagctgaact 1260
caaaaatcaa tggatacaac attgcagatg ccacatttgg tgacacctgt tctactcacc 1320
cccttaagga agagatcatt aaaaagaggt ctgagactgt gaaggctaag atgttgaaac 1380
tggggcctga tggcagaaag gcactctact ctaagccagg tagcaaaaat ggaaggtgga 1440
accctgagac ccacaagttc tgcaaatgtg gggtgagaat acagacatca gcctacactt 1500
gcagcaagtg taggaacaga tcaggagaga ataactcctt tttcaaccac aagcattctg 1560
acatcaccaa gtccaaaatt tcaggcaaga tgaaaggtca aggtccatcg gccgactgat 1620
aagtcgacca ccaccaccac caccactgat aagagctcct cgagcctgca gcagctgaag 1680
ctttggactt ctggcgcgcc caattcgccc tatagtgagt cgtattacgt cgcgctcact 1740
ggccgtcgtt ttacaacgtc gtgactggga aaaccctggc gttacccaac ttaatcgcct 1800
tgcagcacat cccccttcg ccagctggcg taatagcgaa gaggcccgca ccgaaacgcc 1860
cttcccaaca gttgcgcagc ctgaatggcg aatgggagcg ccctgtagcg gcgcattaag 1920

cgcggcgggt gtggtggtta cgcgcagcgt gaccgctaca cttgccagcg ccctagcgcc 1980
cgctcctttc gctttcttcc cttcctttct cgccacgttc gccggctttc cccgtcaagc 2040
tctaaatcgg gggctccctt tagggttccg atttagtgct ttacggcacc tcgaccccaa 2100
aaaacttgat tagggtgatg gttggcctgt agtgggccat agccctgata gacggttttt 2160
cgcccctttga cgttggagtc cacgttcttt aatagtggac tcttgttcca aactggaaca 2220
acactcaacc ctatctcggt ctattctttt gatttataag ggattttgcc gatttcggcc 2280
tattggttaa aaaatgagct gatttaacaa aaatttaacg cgaattttaa caaaatatta 2340
acgcttacaa tttaggtggc acttttcggg gaaatgtgcg cggaacccct atttgtttat 2400
ttttctaaat acattcaaat atgtatccgc tcatgagaca ataaccctga taaatgcttc 2460
aataatattg aaaaaggaag agtatgagta ttcaacattt ccgtgtcgcc cttattccct 2520
tttttgcggc attttgcctt cctgtttttg ctcacccaga aacgctggtg aaagtaaaag 2580
atgctgaaga tcagttgggt gcacgagtgg gttacatcga actggatctc aacagcggta 2640
agatccttga gagttttcgc cccgaagaac gttttccaat gatgagcact tttaaagttc 2700
tgctatgtgg cgcggtatta tcccgtattg acgccgggca agagcaactc ggtcgccgca 2760
tacactattc tcagaatgac ttggttgagt actcaccagt cacagaaaag catcttacgg 2820
atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat aacactgcgg 2880
ccaacttact tctgacaacg atcggaggac cgaaggagct aaccgctttt ttgcacaaca 2940
tgggggatca tgtaactcgc cttgatcgtt gggaaccgga gctgaatgaa gccataccaa 3000
acgacgagcg tgacaccacg atgcctgtag caatggcaac aacgttgcgc aaactattaa 3060
ctggcgaact acttactcta gcttcccggc aacaattaat agactggatg gaggcggata 3120
aagttgcagg accacttctg cgctcggccc ttccggctgg ctggtttatt gctgataaat 3180
ctggagccgg tgagcgtggt tctcgcggta tcattgcagc actggggcca gatggtaagc 3240
cctcccgtat cgtagttatc tacacgacgg ggagtcaggc aactatggat gaacgaaata 3300
gacagatcgc tgagataggt gcctcactga ttaagcattg gtaactgtca gaccaagttt 3360
actcatatat actttagatt gatttaaaac ttcattttta atttaaaagg atctaggtga 3420
agatcctttt tgataatctc atgaccaaaa tcccttaacg tgagttttcg ttccactgag 3480
cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa 3540
tctgctgctt gcaaacaaaa aaaccaccgc taccagcggt ggtttgtttg ccggatcaag 3600
agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata ccaaatactg 3660
ttcttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca ccgcctacat 3720
acctcgctct gctaatcctg ttaccagtgg ctgctgccag tggcgataag tcgtgtctta 3780
ccgggttgga ctcaagacga tagttaccgg ataaggcgca gcggtcgggc tgaacggggg 3840
gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc 3900
gtgagctatg agaaagcgcc acgcttcccg aagggagaaa ggcggacagg tatccggtaa 3960
gcggcagggt cggaacagga gagcgcacga gggagcttcc agggggaaac gcctggtatc 4020
tttatagtcc tgtcgggttt cgccacctct gacttgagcg tcgattttg tgatgctcgt 4080
caggg 4085

<210> 151

<211> 4046
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS9011 (pCLS 7865-cT11 TevD02)

<400> 151
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 60
tggccttttg ctcacatggt ctttcctgcg ttatcccctg attctgtgga taaccgtatt 120
accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca 180
gtgagcgagg aagcggagag cgcccaatac gcaaaccgcc tctccccgcg cgttggccga 240
ttcattaatg cagctggcac gacaggtttc ccgactggaa agcgggcagt gagcgcaacg 300
caattaatgt gagttagctc actcattagg caccccaggc tttacacttt atgcttccgg 360
ctcgtatgtt gtgtggaatt gtgagcggat aacaatttca cacaggaaac agctatgacc 420
atgattacgc caagcgcgtc aattaaccct cactaaaggg aacaaaagct gttaattaag 480
cgcgccacac gcaaacacaa atacacagcg gccttgccac catggtggat ctacgcacgc 540
tcggctacag ccagcagcaa caggagaaga tcaaaccgaa ggttcgttcg acagtggcgc 600
agcaccacga ggcactggtc ggccacgggt ttacacacgc gcacatcgtt gcgttaagcc 660
aacacccggc agcgttaggg accgtcgctg tcaagtatca ggacatgatc gcagcgttgc 720
cagaggcgac acacgaagcg atcgttggcg tcggcaaaca gtggtccggc gcacgcgctc 780
tggaggcctt gctcacggtg gcgggagagt tgagaggtcc accgttacag ttggacacag 840
gccaacttct caagattgca aaacgtggcg gcgtgaccgc agtggaggca gtgcatgcat 900
ggcgcaatgc actgacgggt gccccgctca acttgaccgg agacgcccgg gggatcaggt 960
cacgtgcgtc tcggagcatt gttgcccagt tatctcgccc tgatccggga tccaagtctg 1020
gcatctacca gattaagaac accctgaata acaaagtgta tgttggatct gccaaggact 1080
ttgagaaaag atggaagagg cacttcaagg atctggaaaa gggttgccac agctccatca 1140

agctccaaag gtctttcaac aaacatggca atgtctttga gtgctccatt ttggaagaga 1200
tcccttatga aaaggacctc attatagaga gggaaaactt ctggatcaag gagctgaact 1260
caaaaatcaa tggatacaac attgcagatg ccacatttgg tgacacctgt tctactcacc 1320
cccttaagga agagatcatt aaaaagaggt ctgagactgt gaaggctaag atgttgaaac 1380
tggggcctga tggcagaaag gcactctact ctaagccagg tagcaaaaat ggaaggtgga 1440
accctgagac ccacaagttc tgcaaatgtg gggtgagaat acagacatca gcctacactt 1500
gcagcaagtg taggaacaga tcaggagaga ataactcctt tttcaaccac aagcattctc 1560
aaggtccatc ggccgactga taagtcgacc accaccacca ccaccactga taagagctcc 1620
tcgagcctgc agcagctgaa gctttggact tctggcgcgc ccaattcgcc ctatagtgag 1680
tcgtattacg tcgcgctcac tggccgtcgt tttacaacgt cgtgactggg aaaaccctgg 1740
cgttacccaa cttaatcgcc ttgcagcaca tccccctttc gccagctggc gtaatagcga 1800
agaggcccgc accgaaacgc ccttcccaac agttgcgcag cctgaatggc gaatgggagc 1860
gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac 1920
acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc tcgccacgtt 1980
cgccggcttt ccccgtcaag ctctaaatcg ggggctccct ttagggttcc gatttagtgc 2040
tttacggcac ctcgacccca aaaaacttga ttagggtgat ggttggcctg tagtgggcca 2100
tagccctgat agacggtttt tcgccctttg acgttggagt ccacgttctt taatagtgga 2160
ctcttgttcc aaactggaac aacactcaac cctatctcgg tctattcttt tgatttataa 2220
gggattttgc cgatttcggc ctattggtta aaaaatgagc tgatttaaca aaaatttaac 2280
gcgaatttta acaaaatatt aacgcttaca atttaggtgg cacttttcgg ggaaatgtgc 2340
gcggaacccc tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcatgagac 2400
aataaccctg ataaatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt 2460
tccgtgtcgc ccttattccc ttttttgcgg cattttgcct tcctgttttt gctcacccag 2520
aaacgctggt gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg ggttacatcg 2580
aactggatct caacagcggt aagatccttg agagttttcg ccccgaagaa cgttttccaa 2640
tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt atcccgtatt gacgccgggc 2700
aagagcaact cggtcgccgc atacactatt ctcagaatga cttggttgag tactcaccag 2760
tcacagaaaa gcatcttacg gatggcatga cagtaagaga attatgcagt gctgccataa 2820
ccatgagtga taacactgcg gccaacttac ttctgacaac gatcggagga ccgaaggagc 2880
taaccgcttt tttgcacaac atgggggatc atgtaactcg ccttgatcgt tgggaaccgg 2940
agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgta gcaatggcaa 3000
caacgttgcg caaactatta actggcgaac tacttactct agcttcccgg caacaattaa 3060
tagactggat ggaggcggat aaagttgcag gaccacttct gcgctcggcc cttccggctg 3120
gctggtttat tgctgataaa tctggagccg gtgagcgtgg ttctcgcggt atcattgcag 3180
cactggggcc agatggtaag ccctcccgta tcgtagttat ctacacgacg gggagtcagg 3240
caactatgga tgaacgaaat agacagatcg ctgagatagg tgcctcactg attaagcatt 3300
ggtaactgtc agaccaagtt tactcatata tactttagat tgatttaaaa cttcattttt 3360
aatttaaaag gatctaggtg aagatccttt ttgataatct catgaccaaa atcccttaac 3420
gtgagttttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag 3480
atcctttttt tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg 3540
tggtttgttt gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca 3600
gagcgcagat accaaatact gttcttctag tgtagccgta gttaggccac cacttcaaga 3660
actctgtagc accgcctaca tacctcgctc tgctaatcct gttaccagtg gctgctgcca 3720
gtggcgataa gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc 3780
agcggtcggg ctgaacgggg ggttcgtgca cacagcccag cttggagcga acgacctaca 3840
ccgaactgag atacctacag cgtgagctat gagaaagcgc cacgcttccc gaagggagaa 3900
aggcggacag gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc 3960
caggggaaa cgcctggtat ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc 4020
gtcgattttt gtgatgctcg tcaggg 4046

<210> 152

<211> 598
<212> PRT
<213> Artificial sequence

<220>
<223> RVDs AvrBs3

<400> 152
Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
1 5 10 15

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
20 25 30

```
His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
        35                  40                  45

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
    50                  55                  60

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn
65                  70                  75                  80

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val
                85                  90                  95

Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
            100                 105                 110

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
        115                 120                 125

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
    130                 135                 140

Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala
145                 150                 155                 160

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
                165                 170                 175

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val
            180                 185                 190

Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
        195                 200                 205

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu
    210                 215                 220

Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
225                 230                 235                 240

Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala
                245                 250                 255

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
            260                 265                 270

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
        275                 280                 285

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
    290                 295                 300
```

```
His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
305                 310                 315                 320

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
                325                 330                 335

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn
            340                 345                 350

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val
        355                 360                 365

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
    370                 375                 380

Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu
385                 390                 395                 400

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
                405                 410                 415

Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
            420                 425                 430

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
        435                 440                 445

Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val
    450                 455                 460

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
465                 470                 475                 480

Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu
                485                 490                 495

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
            500                 505                 510

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala
        515                 520                 525

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
    530                 535                 540

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
545                 550                 555                 560

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
                565                 570                 575
```

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
580 585 590

Gly Arg Pro Ala Leu Glu
595

<210> 153

<211> 5798
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS9013

<400> 153
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 60
tggccttttg ctcacatggt ctttcctgcg ttatcccctg attctgtgga taaccgtatt 120
accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca 180
gtgagcgagg aagcggagag cgcccaatac gcaaaccgcc tctccccgcg cgttggccga 240
ttcattaatg cagctggcac gacaggtttc ccgactggaa agcgggcagt gagcgcaacg 300
caattaatgt gagttagctc actcattagg caccccaggc tttacacttt atgcttccgg 360
ctcgtatgtt gtgtggaatt gtgagcggat aacaatttca cacaggaaac agctatgacc 420
atgattacgc caagcgcgtc aattaaccct cactaaaggg aacaaaagct gttaattaag 480
cgcgccacac gcaaacacaa atacacagcg gccttgccac catggtggat ctacgcacgc 540
tcggctacag ccagcagcaa caggagaaga tcaaaccgaa ggttcgttcg acagtggcgc 600
agcaccacga ggcactggtc ggccacgggt ttacacacgc gcacatcgtt gcgttaagcc 660
aacacccggc agcgttaggg accgtcgctg tcaagtatca ggacatgatc gcagcgttgc 720
cagaggcgac acacgaagcg atcgttggcg tcggcaaaca gtggtccggc gcacgcgctc 780
tggaggcctt gctcacggtg gcgggagagt tgagaggtcc accgttacag ttggacacag 840
gccaacttct caagattgca aaacgtggcg gcgtgaccgc agtggaggca gtgcatgcat 900
ggcgcaatgc actgacgggt gccccgctca acttgacccc ggagcaggtg gtggccatcg 960
ccagccacga tggcggcaag caggcgctgg agacggtcca gcggctgttg ccggtgctgt 1020
gccaggccca cggcttgacc ccccagcagg tggtggccat cgccagcaat ggcggtggca 1080
agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc cacggcttga 1140
ccccggagca ggtggtggcc atcgccagca atattggtgg caagcaggcg ctggagacgg 1200
tgcaggcgct gttgccggtg ctgtgccagg cccacggctt gacccccag caggtggtgg 1260
ccatcgccag caatggcggt ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg 1320
tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc agcaatattg 1380
gtggcaagca ggcgctggag acggtgcagg cgctgttgcc ggtgctgtgc caggcccacg 1440
gcttgacccc ggagcaggtg gtggccatcg ccagcaatat tggtggcaag caggcgctgg 1500
agacggtgca ggcgctgttg ccggtgctgt gccaggccca cggcttgacc ccggagcagg 1560
tggtggccat cgccagcaat attggtggca agcaggcgct ggagacggtg caggcgctgt 1620
tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc atcgccagcc 1680
acgatggcgg caagcaggcg ctggagacgg tccagcggct gttgccggtg ctgtgccagg 1740
cccacggctt gaccccggag caggtggtgg ccatcgccag ccacgatggc ggcaagcagg 1800
cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc ttgacccccc 1860
agcaggtggt ggccatcgcc agcaatggcg gtggcaagca ggcgctggag acggtccagc 1920
ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg gtggccatcg 1980
ccagcaatat tggtggcaag caggcgctgg agacggtgca ggcgctgttg ccggtgctgt 2040
gccaggccca cggcttgacc ccggagcagg tggtggccat cgccagcaat attggtggca 2100
agcaggcgct ggagacggtg caggcgctgt tgccggtgct gtgccaggcc cacggcttga 2160
ccccggagca ggtggtggcc atcgccagcc acgatggcgg caagcaggcg ctggagacgg 2220
tccagcggct gttgccggtg ctgtgccagg cccacggctt gaccccggag caggtggtgg 2280
ccatcgccag ccacgatggc ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg 2340
tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc agccacgatg 2400
gcggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc caggcccacg 2460
gcttgacccc ccagcaggtg gtggccatcg ccagcaatgg cggtggcaag caggcgctgg 2520
agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc ccggagcagg 2580
tggtggccat cgccagccac gatggcggca agcaggcgct ggagacggtc cagcggctgt 2640
tgccggtgct gtgccaggcc cacggcttga cccctcagca ggtggtggcc atcgccagca 2700
atggcggcgg caggccggcg ctggagagca ttgttgccca gttatctcgc cctgatccgg 2760
gatccaagtc tggcatctac cagattaaga acaccctgaa taacaaagtg tatgttggat 2820
ctgccaagga ctttgagaaa agatggaaga ggcacttcaa ggatctggaa aagggttgcc 2880
acagctccat caagctccaa aggtctttca acaaacatgg caatgtcttt gagtgctcca 2940

ttttggaaga gatcccttat gaaaaggacc tcattataga gagggaaaac ttctggatca 3000
aggagctgaa ctcaaaaatc aatggataca acattgcaga tgccacattt ggtgacacct 3060
gttctactca ccccccttaag gaagagatca ttaaaaagag gtctgagact gtgaaggcta 3120
agatgttgaa actggggcct gatggcagaa aggcactcta ctctaagcca ggtagcaaaa 3180
atggaaggtg gaaccctgag acccacaagt tctgcaaatg tggggtgaga atacagacat 3240
cagcctacac ttgcagcaag tgtaggaaca gatcaggaga gaataactcc tttttcaacc 3300
acaagcattc tcaaggtcca tcggccgact gataagtcga ccaccaccac caccaccact 3360
gataagagct cctcgagcct gcagcagctg aagctttgga cttctggcgc gcccaattcg 3420
ccctatagtg agtcgtatta cgtcgcgctc actggccgtc gttttacaac gtcgtgactg 3480
ggaaaaccct ggcgttaccc aacttaatcg ccttgcagca catccccctt tcgccagctg 3540
gcgtaatagc gaagaggccc gcaccgaaac gcccttccca acagttgcgc agcctgaatg 3600
gcgaatggga gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg ttacgcgcag 3660
cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct tcccttcctt 3720
tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc ctttagggtt 3780
ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg atggttggcc 3840
tgtagtgggc catagccctg atagacggtt tttcgccctt tgacgttgga gtccacgttc 3900
tttaatagtg gactcttgtt ccaaactgga acaacactca accctatctc ggtctattct 3960
tttgatttat aagggatttt gccgatttcg gcctattggt taaaaaatga gctgatttaa 4020
caaaaattta acgcgaattt taacaaaata ttaacgctta caatttaggt ggcacttttc 4080
ggggaaatgt gcgcggaacc cctatttgtt tatttttcta aatacattca aatatgtatc 4140
cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga 4200
gtattcaaca tttccgtgtc gcccttattc cctttttttgc ggcattttgc cttcctgttt 4260
ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag 4320
tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag 4380
aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgta 4440
ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg 4500
agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca 4560
gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca acgatcggag 4620
gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc 4680
gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg 4740
tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact ctagcttccc 4800
ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg 4860
cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt ggttctcgcg 4920
gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga 4980
cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 5040
tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag attgatttaa 5100
aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat ctcatgacca 5160
aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag 5220
gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac 5280
cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa 5340
ctggcttcag cagagcgcag ataccaaata ctgttcttct agtgtagccg tagttaggcc 5400
accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatc ctgttaccag 5460
tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac 5520
cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttggagc 5580
gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc 5640
ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca 5700
cgagggagct tccagggga aacgcctggt atctttatag tcctgtcggg tttcgccacc 5760
tctgacttga gcgtcgattt ttgtgatgct cgtcaggg 5798

<210> 154

<211> 11148
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS8523

<400> 154
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 60
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 120
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 180
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 240
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 300
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 360
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 420
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 480

ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 540
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 600
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 660
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 720
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 780
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 840
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 900
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 960
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1020
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1080
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1140
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1200
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1260
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1320
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1380
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1440
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1500
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1560
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1620
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1680
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1740
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1800
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1860
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1920
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 1980
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2040
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2100
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2160
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2220
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2280
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2340
gacacgcaaa cacaaataca cagcggcctt gccaccatgg tggatctacg cacgctcggc 2400
tacagccagc agcaacagga gaagatcaaa ccgaaggttc gttcgacagt ggcgcagcac 2460
cacgaggcac tggtcggcca cgggtttaca cacgcgcaca tcgttgcgtt aagccaacac 2520
ccggcagcgt tagggaccgt cgctgtcaag tatcaggaca tgatcgcagc gttgccagag 2580
gcgacacacg aagcgatcgt tggcgtcggc aaacagtggt ccggcgcacg cgctctggag 2640
gccttgctca cggtggcggg agagttgaga ggtccaccgt tacagttgga cacaggccaa 2700
cttctcaaga ttgcaaaacg tggcggcgtg accgcagtgg aggcagtgca tgcatggcgc 2760
aatgcactga cgggtgcccc gctcaacttg accccggagc aggtggtggc catcgccagc 2820
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2880
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 2940
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 3000
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 3060
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 3120
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3180
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 3240
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 3300
accccggagc aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg 3360
gtgcaggcgc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 3420
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 3480
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat 3540
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 3600
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg 3660
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 3720
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 3780
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 3840
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 3900
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag 3960
gcgctggaga cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 4020
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 4080
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 4140
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 4200
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 4260
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 4320
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 4380
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 4440
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 4500
gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg tggccatcgc cagcaatggc 4560

```
ggcggcaggc cggcgctgga gagcattgtt gcccagttat ctcgccctga tccgggatcc   4620
aagtctggca tctaccagat taagaacacc ctgaataaca aagtgtatgt tggatctgcc   4680
aaggactttg agaaaagatg gaagaggcac ttcaaggatc tggaaaaggg ttgccacagc   4740
tccatcaagc tccaaaggtc tttcaacaaa catggcaatg tctttgagtg ctccattttg   4800
gaagagatcc cttatgaaaa ggacctcatt atagagaggg aaaacttctg gatcaaggag   4860
ctgaactcaa aaatcaatgg atacaacatt gcagatgcca catttggtga cacctgttct   4920
actcaccccc ttaaggaaga gatcattaaa aagaggtctg agactgtgaa ggctaagatg   4980
ttgaaactgg ggcctgatgg cagaaaggca ctctactcta agccaggtag caaaaatgga   5040
aggtggaacc ctgagaccca caagttctgc aaatgtgggg tgagaataca gacatcagcc   5100
tacacttgca gcaagtgtag gaacagatca ggagagaata actccttttt caaccacaag   5160
cattctgaca tcaccaagtc caaaatttca ggcaagatga aaggtcaagg tccatcggcc   5220
gactgataac tcgagcgatc ctctagacga gctcctcgag cctgcagcag ctgaagcttt   5280
ggacttcttc gccagaggtt tggtcaagtc tccaatcaag gttgtcggct tgtctacctt   5340
gccagaaatt tacgaaaaga tggaaaaggg tcaaatcgtt ggtagatacg ttgttgacac   5400
ttctaaataa gcgaatttct tatgatttat gatttttatt attaaataag ttataaaaaa   5460
aataagtgta tacaaatttt aaagtgactc ttaggtttta aaacgaaaat tcttattctt   5520
gagtaactct ttcctgtagg tcaggttgct ttctcaggta tagcatgagg tcgctcttat   5580
tgaccacacc tctaccggca tgcaagcttg gcgtaatcat ggtcatagct gtttcctgtg   5640
tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat aaagtgtaaa   5700
gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgctc actgcccgct   5760
ttccagtcgg gaaacctgtc gtgccagcag atctattaca ttatgggtgg tatgttggaa   5820
taaaaatcaa ctatcatcta ctaactagta tttacgttac tagtatatta tcatatacgg   5880
tgttagaaga tgacgcaaat gatgagaaat agtcatctaa attagtggaa gctgaaacgc   5940
aaggattgat aatgtaatag gatcaatgaa tattaacata taaaatgatg ataataatat   6000
ttatagaatt gtgtagaatt gcagattccc ttttatggat tcctaaatcc tcgaggagaa   6060
cttctagtat atctacatac ctaatattat tgccttatta aaaatggaat cccaacaatt   6120
acatcaaaat ccacattctc ttcaaaatca attgtcctgt acttccttgt tcatgtgtgt   6180
tcaaaaacgt tatatttata ggataattat actctatttc tcaacaagta attggttgtt   6240
tggccgagcg gtctaaggcg cctgattcaa gaaatatctt gaccgcagtt aactgtggga   6300
atactcaggt atcgtaagat gcaagagttc gaatctctta gcaaccatta ttttttttcct   6360
caacataacg agaacacaca ggggcgctat cgcacagaat caaattcgat gactggaaat   6420
tttttgttaa tttcagaggt cgcctgacgc atataccttt ttcaactgaa aaattgggag   6480
aaaaaggaaa ggtgagagcc gcggaaccgg cttttcatat agaatagaga agcgttcatg   6540
actaaatgct tgcatcacaa tacttgaagt tgacaatatt atttaaggac ctattgtttt   6600
ttccaatagg tggttagcaa tcgtcttact ttctaacttt tcttaccttt tacatttcag   6660
caatatatat atatatattt caaggatata ccattctaat gtctgcccct aagaagatcg   6720
tcgttttgcc aggtgaccac gttggtcaag aaatcacagc cgaagccatt aaggttctta   6780
aagctatttc tgatgttcgt tccaatgtca agttcgattt cgaaaatcat ttaattggtg   6840
gtgctgctat cgatgctaca ggtgtcccac ttccagatga ggcgctggaa gcctccaaga   6900
aggttgatgc cgttttgtta ggtgctgtgg gtggtcctaa atggggtacc ggtagtgtta   6960
gacctgaaca aggtttacta aaaatccgta aagaacttca attgtacgcc aacttaagac   7020
catgtaactt tgcatccgac tctcttttag acttatctcc aatcaagcca caatttgcta   7080
aaggtactga cttcgttgtt gtcagagaat tagtgggagg tatttacttt ggtaagagaa   7140
aggaagacga tggtgatggt gtcgcttggg atagtgaaca atacaccgtt ccagaagtgc   7200
aaagaatcac aagaatggcc gctttcatgg ccctacaaca tgagccacca ttgcctattt   7260
ggtccttgga taaagctaat gttttggcct cttcaagatt atggagaaaa actgtggagg   7320
aaaccatcaa gaacgaattc cctacattga aggttcaaca tcaattgatt gattctgccg   7380
ccatgatcct agttaagaac ccaacccacc taaatggtat tataatcacc agcaacatgt   7440
ttggtgatat catctccgat gaagcctccg ttatcccagg ttccttgggt ttgttgccat   7500
ctgcgtcctt ggcctctttg ccagacaaga acaccgcatt tggtttgtac gaaccatgcc   7560
acggttctgc tccagatttg ccaaagaata aggtcaaccc tatcgccact atcttgtctg   7620
ctgcaatgat gttgaaattg tcattgaact tgcctgaaga aggtaaggcc attgaagatg   7680
cagttaaaaa ggttttggat gcaggtatca gaactggtga tttaggtggt tccaacagta   7740
ccacggaagt cggtgatgct gtcgccgaag aagttaagaa aatccttgct taaaaagatt   7800
ctctttttt atgatatttg tacataaact ttataaatga aattcataat agaaacgaca   7860
cgaaattaca aaatggaata tgttcatagg gtagacgaaa ctatatacgc aatctacata   7920
catttatcaa gaaggagaaa aaggaggatg taaaggaata caggtaagca aattgatact   7980
aatggctcaa cgtgataagg aaaaagaatt gcactttaac attaatattg acaaggagga   8040
gggcaccaca caaaaagtta ggtgtaacag aaaatcatga aactatgatt cctaatttat   8100
atattggagg attttctcta aaaaaaaaaa aatacaacaa ataaaaaaca ctcaatgacc   8160
tgaccatttg atggagttta agtcaatacc ttcttgaacc atttcccata atggtgaaag   8220
ttccctcaag aattttactc tgtcagaaac ggccttaacg acgtagtcga cctcctcttc   8280
agtactaaat ctaccaatac caaatctgat ggaagaatgg gctaatgcat catccttacc   8340
cagcgcatgt aaaacataag aaggttctag ggaagcagat gtacaggctg aacccgagga   8400
taatgcgata tcccttagtg ccatcaataa agattctcct tccacgtagg cgaaagaaac   8460
gttaacacac cctggataac gatgatctgg agatccgttc aacgtggtat gttcagcgga   8520
taatagacct ttgactaatt tatcggatag tcttttgatg tgagcttggt cgttgtcaaa   8580
ttctttcttc atcaatctcg cagcttcacc aaatcccgct accaatgggg gggccaaagt   8640
```

accagatctg ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg 8700
gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc 8760
ggtatcagca tcgatgaatt ccacggacta tagactatac tagtatactc cgtctactgt 8820
acgatacact tccgctcagg tccttgtcct ttaacgaggc cttaccactc ttttgttact 8880
ctattgatcc agctcagcaa aggcagtgtg atctaagatt ctatcttcgc gatgtagtaa 8940
aactagctag accgagaaag agactagaaa tgcaaaaggc acttctacaa tggctgccat 9000
cattattatc cgatgtgacg ctgcagcttc tcaatgatat tcgaatacgc tttgaggaga 9060
tacagcctaa tatccgacaa actgttttac agatttacga tcgtacttgt tacccatcat 9120
tgaattttga acatccgaac ctgggagttt tccctgaaac agatagtata tttgaacctg 9180
tataataata tatagtctag cgctttacgg aagacaatgt atgtatttcg gttcctggag 9240
aaactattgc atctattgca taggtaatct tgcacgtcgc atccccggtt cattttctgc 9300
gtttccatct tgcacttcaa tagcatatct ttgttaacga agcatctgtg cttcattttg 9360
tagaacaaaa atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt 9420
ttacagaaca gaaatgcaac gcgaaagcgc tattttacca acgaagaatc tgtgcttcat 9480
ttttgtaaaa caaaaatgca acgcgagagc gctaattttt caaacaaaga atctgagctg 9540
cattttaca gaacagaaat gcaacgcgag agcgctattt taccaacaaa gaatctatac 9600
ttctttttg ttctacaaaa atgcatcccg agagcgctat ttttctaaca aagcatctta 9660
gattactttt tttctccttt gtgcgctcta taatgcagtc tcttgataac tttttgcact 9720
gtaggtccgt taaggttaga agaaggctac tttggtgtct attttctctt ccataaaaaa 9780
agcctgactc cacttcccgc gtttactgat tactagcgaa gctgcgggtg cattttttca 9840
agataaaggc atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca 9900
gaaagtgata gcgttgatga ttcttcattg gtcagaaaat tatgaacggt ttcttctatt 9960
ttgtctctat atactacgta taggaaatgt ttacattttc gtattgtttt cgattcactc 10020
tatgaatagt tcttactaca atttttttgt ctaaagagta atactagaga taaacataaa 10080
aaatgtagag gtcgagttta gatgcaagtt caaggagcga aaggtggatg ggtaggttat 10140
atagggatat agcacagaga tatatagcaa agagatactt ttgagcaatg tttgtggaag 10200
cggtattcgc aatattttag tagctcgtta cagtccggtg cgtttttggt ttttgaaag 10260
tgcgtcttca gagcgctttt ggttttcaaa agcgctctga agttcctata ctttctagag 10320
aataggaact tcggaatagg aacttcaaag cgtttccgaa aacgagcgct tccgaaaatg 10380
caacgcgagc tgcgcacata cagctcactg ttcacgtcgc acctatatct gcgtgttgcc 10440
tgtatatata tatacatgag aagaacggca tagtgcgtgt ttatgcttaa atgcgtactt 10500
atatgcgtct atttatgtag gatgaaaggt agtctagtac ctcctgtgat attatcccat 10560
tccatgcggg gtatcgtatg cttccttcag cactaccctt tagctgttct atatgctgcc 10620
actcctcaat tggattagtc tcatccttca atgctatcat ttcctttgat attggatcat 10680
atgcatagta ccgagaaact agtgcgaagt agtgatcagg tattgctgtt atctgatgag 10740
tatacgttgt cctggccacg gcagaagcac gcttatcgct ccaatttccc acaacattag 10800
tcaactccgt taggcccttc attgaaagaa atgaggtcat caaatgtctt ccaatgtgag 10860
attttgggcc atttttata gcaaagattg aataaggcgc atttttcttc aaagctttat 10920
tgtacgatct gactaagtta tcttttaata attggtattc ctgtttattg cttgaagaat 10980
tgccggtcct atttactcgt tttaggactg gttcagaatt catcgatgct cactcaaagg 11040
tcggtaatac ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa 11100
ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttt 11148

<210> 155

<211> 11109
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS8524

<400> 155
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 60
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 120
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 180
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 240
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 300
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 360
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 420
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 480
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 540
tttttgtttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 600
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 660
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 720
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 780
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 840

tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 900
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 960
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1020
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1080
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1140
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1200
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1260
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1320
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1380
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1440
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1500
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1560
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1620
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1680
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1740
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1800
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1860
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1920
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 1980
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2040
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2100
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2160
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2220
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2280
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2340
gacacgcaaa cacaaataca cagcggcctt gccaccatgg tggatctacg cacgctcggc 2400
tacagccagc agcaacagga gaagatcaaa ccgaaggttc gttcgacagt ggcgcagcac 2460
cacgaggcac tggtcggcca cgggtttaca cacgcgcaca tcgttgcgtt aagccaacac 2520
ccggcagcgt tagggaccgt cgctgtcaag tatcaggaca tgatcgcagc gttgccagag 2580
gcgacacacg aagcgatcgt tggcgtcggc aaacagtggt ccggcgcacg cgctctggag 2640
gccttgctca cggtggcggg agagttgaga ggtccaccgt tacagttgga cacaggccaa 2700
cttctcaaga ttgcaaaacg tggcggcgtg accgcagtgg aggcagtgca tgcatggcgc 2760
aatgcactga cgggtgcccc gctcaacttg accccggagc aggtggtggc catcgccagc 2820
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2880
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 2940
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 3000
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 3060
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 3120
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3180
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 3240
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 3300
accccggagc aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg 3360
gtgcaggcgc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 3420
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 3480
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat 3540
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 3600
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg 3660
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 3720
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 3780
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 3840
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 3900
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag 3960
gcgctggaga cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 4020
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 4080
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 4140
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 4200
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 4260
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 4320
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 4380
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 4440
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 4500
gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg tggccatcgc cagcaatggc 4560
ggcggcaggc cggcgctgga gagcattgtt gcccagttat ctcgccctga tccgggatcc 4620
aagtctggca tctaccagat taagaacacc ctgaataaca aagtgtatgt tggatctgcc 4680
aaggactttg agaaaagatg gaagaggcac ttcaaggatc tggaaaaggg ttgccacagc 4740
tccatcaagc tccaaaggtc tttcaacaaa catggcaatg tctttgagtg ctccattttg 4800
gaagagatcc cttatgaaaa ggacctcatt atagagaggg aaaacttctg gatcaaggag 4860
ctgaactcaa aaatcaatgg atacaacatt gcagatgcca catttggtga cacctgttct 4920

actcaccccc ttaaggaaga gatcattaaa aagaggtctg agactgtgaa ggctaagatg 4980
ttgaaactgg ggcctgatgg cagaaaggca ctctactcta agccaggtag caaaaatgga 5040
aggtggaacc ctgagaccca caagttctgc aaatgtgggg tgagaataca gacatcagcc 5100
tacacttgca gcaagtgtag gaacagatca ggagagaata actccttttt caaccacaag 5160
cattctcaag gtccatcggc cgactgataa ctcgagcgat cctctagacg agctcctcga 5220
gcctgcagca gctgaagctt tggacttctt cgccagaggt ttggtcaagt ctccaatcaa 5280
ggttgtcggc ttgtctacct tgccagaaat ttacgaaaag atggaaaagg gtcaaatcgt 5340
tggtagatac gttgttgaca cttctaaata agcgaatttc ttatgattta tgatttttat 5400
tattaaataa gttataaaaa aaataagtgt atacaaattt taaagtgact cttaggtttt 5460
aaaacgaaaa ttcttattct tgagtaactc tttcctgtag gtcaggttgc tttctcaggt 5520
atagcatgag gtcgctctta ttgaccacac ctctaccggc atgcaagctt ggcgtaatca 5580
tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga 5640
gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact cacattaatt 5700
gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagca gatctattac 5760
attatgggtg gtatgttgga ataaaaatca actatcatct actaactagt atttacgtta 5820
ctagtatatt atcatatacg gtgttagaag atgacgcaaa tgatgagaaa tagtcatcta 5880
aattagtgga agctgaaacg caaggattga taatgtaata ggatcaatga atattaacat 5940
ataaaatgat gataataata tttatagaat tgtgtagaat tgcagattcc cttttatgga 6000
ttcctaaatc ctcgaggaga acttctagta tatctacata cctaatatta ttgccttatt 6060
aaaaatggaa tcccaacaat tacatcaaaa tccacattct cttcaaaatc aattgtcctg 6120
tacttccttg ttcatgtgtg ttcaaaaacg ttatatttat aggataatta tactctattt 6180
ctcaacaagt aattggttgt ttggccgagc ggtctaaggc gcctgattca agaaatatct 6240
tgaccgcagt taactgtggg aatactcagg tatcgtaaga tgcaagagtt cgaatctctt 6300
agcaaccatt atttttttcc tcaacataac gagaacacac aggggcgcta tcgcacagaa 6360
tcaaattcga tgactggaaa ttttttgtta atttcagagg tcgcctgacg catatacctt 6420
tttcaactga aaaattggga gaaaaaggaa aggtgagagc cgcggaaccg gcttttcata 6480
tagaatagag aagcgttcat gactaaatgc ttgcatcaca atacttgaag ttgacaatat 6540
tatttaagga cctattgttt tttccaatag gtggttagca atcgtcttac tttctaactt 6600
ttcttacctt ttacatttca gcaatatata tatatatatt tcaaggatat accattctaa 6660
tgtctgcccc taagaagatc gtcgttttgc caggtgacca cgttggtcaa gaaatcacag 6720
ccgaagccat taaggttctt aaagctattt ctgatgttcg ttccaatgtc aagttcgatt 6780
tcgaaaatca tttaattggt ggtgctgcta tcgatgctac aggtgtccca cttccagatg 6840
aggcgctgga agcctccaag aaggttgatg ccgttttgtt aggtgctgtg ggtggtccta 6900
aatggggtac cggtagtgtt agacctgaac aaggtttact aaaaatccgt aaagaacttc 6960
aattgtacgc caacttaaga ccatgtaact ttgcatccga ctctctttta gacttatctc 7020
caatcaagcc acaatttgct aaaggtactg acttcgttgt tgtcagagaa ttagtgggag 7080
gtatttactt tggtaagaga aaggaagacg atggtgatgg tgtcgcttgg gatagtgaac 7140
aatacaccgt tccagaagtg caaagaatca caagaatggc cgctttcatg gccctacaac 7200
atgagccacc attgcctatt tggtccttgg ataaagctaa tgttttggcc tcttcaagat 7260
tatggagaaa aactgtggag gaaaccatca agaacgaatt ccctacattg aaggttcaac 7320
atcaattgat tgattctgcc gccatgatcc tagttaagaa cccaacccac ctaaatggta 7380
ttataatcac cagcaacatg tttggtgata tcatctccga tgaagcctcc gttatcccag 7440
gttccttggg tttgttgcca tctgcgtcct tggcctcttt gccagacaag aacaccgcat 7500
ttggtttgta cgaaccatgc cacggttctg ctccagattt gccaaagaat aaggtcaacc 7560
ctatcgccac tatcttgtct gctgcaatga tgttgaaatt gtcattgaac ttgcctgaag 7620
aaggtaaggc cattgaagat gcagttaaaa aggttttgga tgcaggtatc agaactggtg 7680
atttaggtgg ttccaacagt accacggaag tcggtgatgc tgtcgccgaa gaagttaaga 7740
aaatccttgc ttaaaaagat tctctttttt tatgatattt gtacataaac tttataaatg 7800
aaattcataa tagaaacgac acgaaattac aaaatggaat atgttcatag ggtagacgaa 7860
actatatacg caatctacat acatttatca agaaggagaa aaaggaggat gtaaaggaat 7920
acaggtaagc aaattgatac taatggctca acgtgataag gaaaaagaat tgcactttaa 7980
cattaatatt gacaaggagg agggcaccac acaaaaagtt aggtgtaaca gaaaatcatg 8040
aaactatgat tcctaattta tatattggag gattttctct aaaaaaaaaa aaatacaaca 8100
aataaaaaac actcaatgac ctgaccattt gatggagttt aagtcaatac cttcttgaac 8160
catttcccat aatggtgaaa gttccctcaa gaattttact ctgtcagaaa cggccttaac 8220
gacgtagtcg acctcctctt cagtactaaa tctaccaata ccaaatctga tggaagaatg 8280
ggctaatgca tcatccttac ccagcgcatg taaaacataa gaaggttcta gggaagcaga 8340
tgtacaggct gaacccgagg ataatgcgat atcccttagt gccatcaata aagattctcc 8400
ttccacgtag gcgaaagaaa cgttaacaca ccctggataa cgatgatctg gagatccgtt 8460
caacgtggta tgttcagcgg ataatagacc tttgactaat ttatcggata gtcttttgat 8520
gtgagcttgg tcgttgtcaa attctttctt catcaatctc gcagcttcac caaatcccgc 8580
taccaatggg ggggccaaag taccagatct gctgcattaa tgaatcggcc aacgcgcggg 8640
gagaggcggt ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc 8700
ggtcgttcgg ctgcggcgag cggtatcagc atcgatgaat tccacggact atagactata 8760
ctagtatact ccgtctactg tacgatacac ttccgctcag gtccttgtcc tttaacgagg 8820
ccttaccact cttttgttac tctattgatc cagctcagca aaggcagtgt gatctaagat 8880
tctatcttcg cgatgtagta aaactagcta gaccgagaaa gagactagaa atgcaaaagg 8940
cacttctaca atggctgcca tcattattat ccgatgtgac gctgcagctt ctcaatgata 9000

ttcgaatacg ctttgaggag atacagccta atatccgaca aactgtttta cagatttacg 9060
atcgtacttg ttacccatca ttgaattttg aacatccgaa cctgggagtt ttccctgaaa 9120
cagatagtat atttgaacct gtataataat atatagtcta gcgctttacg gaagacaatg 9180
tatgtatttc ggttcctgga gaaactattg catctattgc ataggtaatc ttgcacgtcg 9240
catccccggt tcatttttctg cgtttccatc ttgcacttca atagcatatc tttgttaacg 9300
aagcatctgt gcttcatttt gtagaacaaa aatgcaacgc gagagcgcta atttttcaaa 9360
caaagaatct gagctgcatt tttacagaac agaaatgcaa cgcgaaagcg ctattttacc 9420
aacgaagaat ctgtgcttca tttttgtaaa acaaaaatgc aacgcgagag cgctaatttt 9480
tcaaacaaag aatctgagct gcatttttac agaacagaaa tgcaacgcga gagcgctatt 9540
ttaccaacaa agaatctata cttctttttt gttctacaaa aatgcatccc gagagcgcta 9600
tttttctaac aaagcatctt agattacttt ttttctcctt tgtgcgctct ataatgcagt 9660
ctcttgataa ctttttgcac tgtaggtccg ttaaggttag aagaaggcta ctttggtgtc 9720
tattttctct tccataaaaa aagcctgact ccacttcccg cgtttactga ttactagcga 9780
agctgcgggt gcatttttttc aagataaagg catccccgat tatattctat accgatgtgg 9840
attgcgcata ctttgtgaac agaaagtgat agcgttgatg attcttcatt ggtcagaaaa 9900
ttatgaacgg tttcttctat tttgtctcta tatactacgt ataggaaatg tttacatttt 9960
cgtattgttt tcgattcact ctatgaatag ttcttactac aatttttttg tctaaagagt 10020
aatactagag ataaacataa aaaatgtaga ggtcgagttt agatgcaagt tcaaggagcg 10080
aaaggtggat gggtaggtta tatagggata tagcacagag atatatagca aagagatact 10140
tttgagcaat gtttgtggaa gcggtattcg caatatttta gtagctcgtt acagtccggt 10200
gcgtttttgg ttttttgaaa gtgcgtcttc agagcgcttt tggttttcaa aagcgctctg 10260
aagttcctat actttctaga gaataggaac ttcggaatag gaacttcaaa gcgtttccga 10320
aaacgagcgc ttccgaaaat gcaacgcgag ctgcgcacat acagctcact gttcacgtcg 10380
cacctatatc tgcgtgttgc ctgtatatat atatacatga gaagaacggc atagtgcgtg 10440
tttatgctta aatgcgtact tatatgcgtc tatttatgta ggatgaaagg tagtctagta 10500
cctcctgtga tattatccca ttccatgcgg ggtatcgtat gcttccttca gcactaccct 10560
ttagctgttc tatatgctgc cactcctcaa ttggattagt ctcatccttc aatgctatca 10620
tttcctttga tattggatca tatgcatagt accgagaaac tagtgcgaag tagtgatcag 10680
gtattgctgt tatctgatga gtatacgttg tcctggccac ggcagaagca cgcttatcgc 10740
tccaatttcc cacaacatta gtcaactccg ttaggccctt cattgaaaga aatgaggtca 10800
tcaaatgtct tccaatgtga gattttgggc catttttttat agcaaagatt gaataaggcg 10860
catttttctt caaagcttta ttgtacgatc tgactaagtt atcttttaat aattggtatt 10920
cctgtttatt gcttgaagaa ttgccggtcc tatttactcg ttttaggact ggttcagaat 10980
tcatcgatgc tcactcaaag gtcggtaata cggttatcca cagaatcagg ggataacgca 11040
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 11100
ctggcgttt 11109

<210> 156

<211> 8334
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS0542

<400> 156
tgtgtgcggc cgactgataa ctcgagcgat cctctagacg agctcctcga gcctgcagca 60
gctgaagctt tggacttctt cgccagaggt ttggtcaagt ctccaatcaa ggttgtcggc 120
ttgtctacct tgccagaaat ttacgaaaag atggaaaagg gtcaaatcgt tggtagatac 180
gttgttgaca cttctaaata agcgaatttc ttatgattta tgatttttat tattaaataa 240
gttataaaaa aaataagtgt atacaaattt taaagtgact cttaggtttt aaaacgaaaa 300
ttcttattct tgagtaactc tttcctgtag gtcaggttgc tttctcaggt atagcatgag 360
gtcgctctta ttgaccacac ctctaccggc atgcaagctt ggcgtaatca tggtcatagc 420
tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca 480
taaagtgtaa agcctggggt gcctaatgag tgagctaact cacattaatt gcgttgcgct 540
cactgcccgc tttccagtcg ggaaacctgt cgtgccagca gatctattac attatgggtg 600
gtatgttgga ataaaaatca actatcatct actaactagt atttacgtta ctagtatatt 660
atcatatacg gtgttagaag atgacgcaaa tgatgagaaa tagtcatcta aattagtgga 720
agctgaaacg caaggattga taatgtaata ggatcaatga atattaacat ataaaatgat 780
gataataata tttatagaat tgtgtagaat tgcagattcc cttttatgga ttcctaaatc 840
ctcgaggaga acttctagta tatctacata cctaatatta ttgccttatt aaaaatggaa 900
tcccaacaat tacatcaaaa tccacattct cttcaaaatc aattgtcctg tacttccttg 960
ttcatgtgtg ttcaaaaacg ttatatttat aggataatta tactctattt ctcaacaagt 1020
aattggttgt ttggccgagc ggtctaaggc gcctgattca agaaatatct tgaccgcagt 1080
taactgtggg aatactcagg tatcgtaaga tgcaagagtt cgaatctctt agcaaccatt 1140
atttttttcc tcaacataac gagaacacac aggggcgcta tcgcacagaa tcaaattcga 1200

```
tgactggaaa ttttttgtta atttcagagg tcgcctgacg catatacctt tttcaactga   1260
aaaattggga gaaaaaggaa aggtgagagc cgcggaaccg gcttttcata tagaatagag   1320
aagcgttcat gactaaatgc ttgcatcaca atacttgaag ttgacaatat tatttaagga   1380
cctattgttt tttccaatag gtggttagca atcgtcttac tttctaactt ttcttacctt   1440
ttacatttca gcaatatata tatatatatt tcaaggatat accattctaa tgtctgcccc   1500
taagaagatc gtcgttttgc caggtgacca cgttggtcaa gaaatcacag ccgaagccat   1560
taaggttctt aaagctattt ctgatgttcg ttccaatgtc aagttcgatt tcgaaaatca   1620
tttaattggt ggtgctgcta tcgatgctac aggtgtccca cttccagatg aggcgctgga   1680
agcctccaag aaggttgatg ccgttttgtt aggtgctgtg ggtggtccta aatggggtac   1740
cggtagtgtt agacctgaac aaggtttact aaaaatccgt aaagaacttc aattgtacgc   1800
caacttaaga ccatgtaact ttgcatccga ctctctttta gacttatctc caatcaagcc   1860
acaatttgct aaaggtactg acttcgttgt tgtcagagaa ttagtgggag gtatttactt   1920
tggtaagaga aaggaagacg atggtgatgg tgtcgcttgg gatagtgaac aatacaccgt   1980
tccagaagtg caaagaatca caagaatggc cgctttcatg gccctacaac atgagccacc   2040
attgcctatt tggtccttgg ataaagctaa tgttttggcc tcttcaagat tatggagaaa   2100
aactgtggag gaaaccatca agaacgaatt ccctacattg aaggttcaac atcaattgat   2160
tgattctgcc gccatgatcc tagttaagaa cccaacccac ctaaatggta ttataatcac   2220
cagcaacatg tttggtgata tcatctccga tgaagcctcc gttatcccag gttccttggg   2280
tttgttgcca tctgcgtcct tggcctcttt gccagacaag aacaccgcat ttggtttgta   2340
cgaaccatgc cacggttctg ctccagattt gccaaagaat aaggtcaacc ctatcgccac   2400
tatcttgtct gctgcaatga tgttgaaatt gtcattgaac ttgcctgaag aaggtaaggc   2460
cattgaagat gcagttaaaa aggttttgga tgcaggtatc agaactggtg atttaggtgg   2520
ttccaacagt accacggaag tcggtgatgc tgtcgccgaa gaagttaaga aaatccttgc   2580
ttaaaaagat tctctttttt tatgatattt gtacataaac tttataaatg aaattcataa   2640
tagaaacgac acgaaattac aaaatggaat atgttcatag ggtagacgaa actatatacg   2700
caatctacat acatttatca agaaggagaa aaaggaggat gtaaaggaat acaggtaagc   2760
aaattgatac taatggctca acgtgataag gaaaaagaat tgcactttaa cattaatatt   2820
gacaaggagg agggcaccac acaaaaagtt aggtgtaaca gaaaatcatg aaactatgat   2880
tcctaattta tatattggag gattttctct aaaaaaaaaa aaatacaaca aataaaaaac   2940
actcaatgac ctgaccattt gatggagttt aagtcaatac cttcttgaac catttcccat   3000
aatggtgaaa gttccctcaa gaattttact ctgtcagaaa cggccttaac gacgtagtcg   3060
acctcctctt cagtactaaa tctaccaata ccaaatctga tggaagaatg ggctaatgca   3120
tcatccttac ccagcgcatg taaaacataa gaaggttcta gggaagcaga tgtacaggct   3180
gaacccgagg ataatgcgat atcccttagt gccatcaata aagattctcc ttccacgtag   3240
gcgaaagaaa cgttaacaca ccctggataa cgatgatctg gagatccgtt caacgtggta   3300
tgttcagcgg ataatagacc tttgactaat ttatcggata gtcttttgat gtgagcttgg   3360
tcgttgtcaa attctttctt catcaatctc gcagcttcac caaatcccgc taccaatggg   3420
ggggccaaag taccagatct gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt   3480
ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg   3540
ctgcggcgag cggtatcagc atcgatgaat tccacggact atagactata ctagtatact   3600
ccgtctactg tacgatacac ttccgctcag gtccttgtcc tttaacgagg ccttaccact   3660
cttttgttac tctattgatc cagctcagca aaggcagtgt gatctaagat tctatcttcg   3720
cgatgtagta aaactagcta gaccgagaaa gagactagaa atgcaaaagg cacttctaca   3780
atggctgcca tcattattat ccgatgtgac gctgcagctt ctcaatgata ttcgaatacg   3840
ctttgaggag atacagccta atatccgaca aactgtttta cagatttacg atcgtacttg   3900
ttacccatca ttgaattttg aacatccgaa cctgggagtt ttccctgaaa cagatagtat   3960
atttgaacct gtataataat atatagtcta gcgctttacg gaagacaatg tatgtatttc   4020
ggttcctgga gaaactattg catctattgc ataggtaatc ttgcacgtcg catccccggt   4080
tcattttctg cgtttccatc ttgcacttca atagcatatc tttgttaacg aagcatctgt   4140
gcttcatttt gtagaacaaa aatgcaacgc gagagcgcta atttttcaaa caaagaatct   4200
gagctgcatt tttacagaac agaaatgcaa cgcgaaagcg ctattttacc aacgaagaat   4260
ctgtgcttca tttttgtaaa acaaaaatgc aacgcgagag cgctaatttt tcaaacaaag   4320
aatctgagct gcatttttac agaacagaaa tgcaacgcga gagcgctatt ttaccaacaa   4380
agaatctata cttctttttt gttctacaaa aatgcatccc gagagcgcta tttttctaac   4440
aaagcatctt agattacttt ttttctcctt tgtgcgctct ataatgcagt ctcttgataa   4500
ctttttgcac tgtaggtccg ttaaggttag aagaaggcta ctttggtgtc tattttctct   4560
tccataaaaa aagcctgact ccacttcccg cgtttactga ttactagcga agctgcgggt   4620
gcatttttttc aagataaagg catccccgat tatattctat accgatgtgg attgcgcata   4680
ctttgtgaac agaaagtgat agcgttgatg attcttcatt ggtcagaaaa ttatgaacgg   4740
tttcttctat tttgtctcta tatactacgt ataggaaatg tttacatttt cgtattgttt   4800
tcgattcact ctatgaatag ttcttactac aatttttttg tctaaagagt aatactagag   4860
ataaacataa aaaatgtaga ggtcgagttt agatgcaagt tcaaggagcg aaaggtggat   4920
gggtaggtta tatagggata tagcacagag atatatagca aagagatact tttgagcaat   4980
gtttgtggaa gcggtattcg caatatttta gtagctcgtt acagtccggt gcgtttttgg   5040
ttttttgaaa gtgcgtcttc agagcgcttt tggttttcaa aagcgctctg aagttcctat   5100
actttctaga gaataggaac ttcggaatag gaacttcaaa gcgtttccga aaacgagcgc   5160
ttccgaaaat gcaacgcgag ctgcgcacat acagctcact gttcacgtcg cacctatatc   5220
tgcgtgttgc ctgtatatat atatacatga gaagaacggc atagtgcgtg tttatgctta   5280
```

aatgcgtact tatatgcgtc tatttatgta ggatgaaagg tagtctagta cctcctgtga 5340
tattatccca ttccatgcgg ggtatcgtat gcttccttca gcactaccct ttagctgttc 5400
tatatgctgc cactcctcaa ttggattagt ctcatccttc aatgctatca tttcctttga 5460
tattggatca tatgcatagt accgagaaac tagtgcgaag tagtgatcag gtattgctgt 5520
tatctgatga gtatacgttg tcctggccac ggcagaagca cgcttatcgc tccaatttcc 5580
cacaacatta gtcaactccg ttaggccctt cattgaaaga aatgaggtca tcaaatgtct 5640
tccaatgtga gattttgggc catttttat agcaaagatt gaataaggcg catttttctt 5700
caaagcttta ttgtacgatc tgactaagtt atcttttaat aattggtatt cctgtttatt 5760
gcttgaagaa ttgccggtcc tatttactcg ttttaggact ggttcagaat tcatcgatgc 5820
tcactcaaag gtcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 5880
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt 5940
tccataggct ccgccccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc 6000
gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct 6060
ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg 6120
tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca 6180
agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta tccggtaact 6240
atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta 6300
acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtggccta 6360
actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag ccagttacct 6420
tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt agcggtggtt 6480
tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga 6540
tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttggtca 6600
tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat 6660
caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg 6720
cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc cccgtcgtgt 6780
agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag 6840
acccacgctc accggctcca gatttatcag caataaacca gccagccgga agggccgagc 6900
gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag 6960
ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca 7020
tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa 7080
ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc ggtcctccga 7140
tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca gcactgcata 7200
attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca 7260
agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 7320
ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa cgttcttcgg 7380
ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg 7440
cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag 7500
gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatac 7560
tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggataca 7620
tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag 7680
tgccacctga cgtccgatca aaaatcatcg cttcgctgat taattacccc agaaataagg 7740
ctaaaaaact aatcgcatta tcatcctatg gttgttaatt tgattcgttc atttgaaggt 7800
ttgtggggcc aggttactgc caatttttcc tcttcataac cataaaagct agtattgtag 7860
aatctttatt gttcggagca gtgcggcgcg aggcacatct gcgtttcagg aacgcgaccg 7920
gtgaagacga ggacgcacgg aggagagtct tccttcggag ggctgtcacc cgctcggcgg 7980
cttctaatcc gtacttcaat atagcaatga gcagttaagc gtattactga aagttccaaa 8040
gagaaggttt ttttaggcta atcgacctcg agcagatccg ccaggcgtgt atatagcgtg 8100
gatggccagg caactttagt gctgacacat acaggcatat atatatgtgt gcgacgacac 8160
atgatcatat ggcatgcatg tgctctgtat gtatataaaa ctcttgtttt cttcttttct 8220
ctaaatattc tttccttata cattaggtcc tttgtagcat aaattactat acttctatag 8280
acacgcaaac acaaatacac agcggccttg ccaccatggc cggcgcgccc acta 8334

<210> 157

<211> 43
<212> DNA
<213> Artificial sequence

<220>
<223> Avr05

<400> 157
tatataaacc taaccctcta ggtaagaggg ttaggtttat ata 43

<210> 158

<211> 44
<212> DNA
<213> Artificial sequence

<220>
<223> Avr06

<400> 158
tatataaacc taaccctcta aggtaagagg gttaggttta tata 44

<210> 159

<211> 45
<212> DNA
<213> Artificial sequence

<220>
<223> Avr07

<400> 159
tatataaacc taaccctcta aggtacagag ggttaggttt atata 45

<210> 160

<211> 46
<212> DNA
<213> Artificial sequence

<220>
<223> Avr08

<400> 160
tatataaacc taaccctctg aaggtacaga gggttaggtt tatata 46

<210> 161

<211> 47
<212> DNA
<213> Artificial sequence

<220>
<223> Avr09

<400> 161
tatataaacc taaccctctg aaggtaccag agggttaggt ttatata 47

<210> 162

<211> 48
<212> DNA
<213> Artificial sequence

<220>
<223> Avr10

<400> 162
tatataaacc taaccctctt gaaggtacca gagggttagg tttatata 48

<210> 163

<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Avr11

<400> 163
tatataaacc taaccctctt gaaggtacct agagggttag gtttatata 49

<210> 164

<211> 50
<212> DNA
<213> Artificial sequence

<220>
<223> Avr12

<400> 164
tatataaacc taaccctcta tgaaggtacc tagagggtta ggtttatata 50

<210> 165

<211> 51
<212> DNA
<213> Artificial sequence

<220>
<223> Avr13

<400> 165
tatataaacc taaccctcta tgaaggtacc ttagagggtt aggtttatat a 51

<210> 166

<211> 52
<212> DNA
<213> Artificial sequence

<220>
<223> Avr14

<400> 166
tatataaacc taaccctctc atgaaggtac cttagagggt taggtttata ta 52

<210> 167

<211> 53
<212> DNA
<213> Artificial sequence

<220>
<223> AVR15

<400> 167
tatataaacc taaccctctt agcatgaagg taccagaggg ttaggtttat ata 53

<210> 168

<211> 54
<212> DNA
<213> Artificial sequence

<220>
<223> Avr16

<400> 168
tatataaacc taaccctctg catgaaggta ccttgagagg gttaggttta tata 54

<210> 169

<211> 55
<212> DNA
<213> Artificial sequence

<220>
<223> Avr17

<400> 169
tatataaacc taaccctctg catgaaggta ccttgtagag ggttaggttt atata 55

<210> 170

<211> 56
<212> DNA
<213> Artificial sequence

<220>
<223> Avr18

<400> 170
tatataaacc taaccctcta gcatgaaggt accttgtaga gggttaggtt tatata 56

<210> 171

<211> 57
<212> DNA
<213> Artificial sequence

<220>
<223> Avr19

<400> 171
tatataaacc taaccctcta gcatgaaggt accttgtcag agggttaggt ttatata 57

<210> 172

<211> 58
<212> DNA
<213> Artificial sequence

<220>
<223> Avr20

<400> 172
tatataaacc taaccctctt agcatgaagg taccttgtca gagggttagg tttatata 58

<210> 173

<211> 59
<212> DNA
<213> Artificial sequence

<220>
<223> Avr21

<400> 173
tatataaacc taaccctctt agcatgaagg taccttgtcg agagggttag gtttatata 59

<210> 174

<211> 60
<212> DNA
<213> Artificial sequence

<220>
<223> Avr22

<400> 174
tatataaacc taaccctctt agcatgaagg taccttgtcg tagagggtta ggtttatata 60

<210> 175

<211> 61
<212> DNA
<213> Artificial sequence

<220>
<223> Avr23

<400> 175
tatataaacc taaccctctc tagcatgaag gtaccttgtc gtagagggtt aggtttatat 60
a 61

<210> 176

<211> 62
<212> DNA
<213> Artificial sequence

<220>
<223> Avr24

<400> 176
tatataaacc taaccctctc tagcatgaag gtaccttgtc gttagagggt taggtttata 60
ta 62

<210> 177

<211> 63
<212> DNA
<213> Artificial sequence

<220>
<223> Avr25

<400> 177
tatataaacc taaccctcta ctagcatgaa ggtaccttgt cgttagaggg ttaggtttat 60
ata 63

<210> 178

<211> 64
<212> DNA
<213> Artificial sequence

<220>
<223> Avr26

<400> 178
tatataaacc taaccctcta ctagcatgaa ggtaccttgt cgttgagagg gttaggttta 60

tata 64

<210> 179

<211> 65
<212> DNA
<213> Artificial sequence

<220>
<223> Avr27

<400> 179
tatataaacc taaccctctc actagcatga aggtaccttg tcgttgagag ggttaggttt 60
atata 65

<210> 180

<211> 66
<212> DNA
<213> Artificial sequence

<220>
<223> Avr28

<400> 180
tatataaacc taaccctctc actagcatga aggtaccttg tcgttgaaga gggttaggtt 60
tatata 66

<210> 181

<211> 67
<212> DNA
<213> Artificial sequence

<220>
<223> Avr29

<400> 181
tatataaacc taaccctctc cactagcatg aaggtacctt gtcgttgaag agggttaggt 60
ttatata 67

<210> 182

<211> 68
<212> DNA
<213> Artificial sequence

<220>
<223> Avr30

<400> 182
tatataaacc taaccctctc cactagcatg aaggtacctt gtcgttgata gagggttagg 60
tttatata 68

<210> 183

<211> 69
<212> DNA
<213> Artificial sequence

<220>
<223> Avr31

<400> 183
tatataaacc taaccctcta ccactagcat gaaggtacct tgtcgttgat agagggttag 60
gtttatata 69

<210> 184

<211> 70
<212> DNA
<213> Artificial sequence

<220>
<223> Avr32

<400> 184
tatataaacc taaccctcta ccactagcat gaaggtacct tgtcgttgat tagagggtta 60
ggtttatata 70

<210> 185

<211> 71
<212> DNA
<213> Artificial sequence

<220>
<223> Avr33

<400> 185
tatataaacc taaccctctg accactagca tgaaggtacc ttgtcgttga ttagagggtt 60
aggtttatat a 71

<210> 186

<211> 72
<212> DNA
<213> Artificial sequence

<220>
<223> Avr34

<400> 186
tatataaacc taaccctctg accactagca tgaaggtacc ttgtcgttga ttcagagggt 60
taggtttata ta 72

<210> 187

<211> 73
<212> DNA
<213> Artificial sequence

<220>
<223> Avr35

<400> 187
tatataaacc taaccctctt gaccactagc atgaaggtac cttgtcgttg attcagaggg 60
ttaggtttat ata 73

<210> 188

<211> 74
<212> DNA
<213> Artificial sequence

<220>

<223> Avr36

<400> 188
tatataaacc taaccctctt gaccactagc atgaaggtac cttgtcgttg attcaagagg 60
gttaggttta tata 74

<210> 189

<211> 75
<212> DNA
<213> Artificial sequence

<220>
<223> Avr37

<400> 189
tatataaacc taaccctctc tgaccactag catgaaggta ccttgtcgtt gattcaagag 60
ggttaggttt atata 75

<210> 190

<211> 76
<212> DNA
<213> Artificial sequence

<220>
<223> Avr38

<400> 190
tatataaacc taaccctctc tgaccactag catgaaggta ccttgtcgtt gattcagaga 60
gggttaggtt tatata 76

<210> 191

<211> 77
<212> DNA
<213> Artificial sequence

<220>
<223> Avr39

<400> 191
tatataaacc taaccctctt ctgaccacta gcatgaaggt accttgtcgt tgattcagag 60
agggttaggt ttatata 77

<210> 192

<211> 78
<212> DNA
<213> Artificial sequence

<220>
<223> Avr40

<400> 192
tatataaacc taaccctctt ctgaccacta gcatgaaggt accttgtcgt tgattcagta 60
gagggttagg tttatata 78

<210> 193

<211> 5672
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS1853

<400> 193
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg gcttatcgaa atgaattcga ctcactgttg ggagacccaa gctggctagt 900
taagctatca caagtttgta caaaaaagca ggctggcgcg cctacacagc ggccttgcca 960
ccatggccaa taccaaatat aacaaagagt tcctgctgta cctggccggc tttgtggacg 1020
gtgacggtag catcatcgct cagattaaac caaaccagtc ttataagttt aaacatcagc 1080
taagcttgac ctttcaggtg actcaaaaga cccagcgccg ttggtttctg gacaaactag 1140
tggatgaaat tggcgttggt tacgtacgtg atcgcggatc cgtttccaac tacatcttaa 1200
gcgaaatcaa gccgctgcac aacttcctga ctcaactgca gccgtttctg aaactgaaac 1260
agaaacaggc aaacctggtt ctgaaaatta tcgaacagct gccgtctgca aaagaatccc 1320
cggacaaatt cctggaagtt tgtacctggg tggatcagat tgcagctctg aacgattcta 1380
agacgcgtaa aaccacttct gaaaccgttc gtgctgtgct ggacagcctg agcgagaaga 1440
agaaatcctc cccggcggcc gactgataac tcgagcgcta gcacccagct ttcttgtaca 1500
aagtggtgat ctagagggcc cgcggttcga aggtaagcct atccctaacc ctctcctcgg 1560
tctcgattct acgcgtaccg gttagtaatg agtttaaacg ggggaggcta actgaaacac 1620
ggaaggagac aataccggaa ggaacccgcg ctatgacggc aataaaaaga cagaataaaa 1680
cgcacgggtg ttgggtcgtt tgttcataaa cgcggggttc ggtcccaggg ctggcactct 1740
gtcgataccc caccgagacc ccattggggc caatacgccc gcgtttcttc cttttcccca 1800
ccccaccccc caagttcggg tgaaggccca gggctcgcag ccaacgtcgg ggcggcaggc 1860
cctgccatag cagatctgcg cagctggggc tctagggggt atccccacgc gccctgtagc 1920
ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac acttgccagc 1980
gccctagcgc ccgctccttt cgctttcttc ccttcctttc tcgccacgtt cgccggcttt 2040
ccccgtcaag ctctaaatcg gggcatccct ttagggttcc gatttagtgc tttacggcac 2100
ctcgacccca aaaaacttga ttagggtgat ggttcacgta gtgggccatc gccctgatag 2160
acggtttttc gccctttgac gttggagtcc acgttcttta atagtggact cttgttccaa 2220
actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg gattttgggg 2280
atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc gaattaattc 2340
tgtggaatgt gtgtcagtta gggtgtggaa agtccccagg ctccccagca ggcagaagta 2400
tgcaaagcat gcatctcaat tagtcagcaa ccaggtgtgg aaagtcccca ggctccccag 2460
caggcagaag tatgcaaagc atgcatctca attagtcagc aaccatagtc ccgccctaa 2520
ctccgcccat cccgccccta actccgccca gttccgccca ttctccgccc catggctgac 2580
taatttttt tatttatgca gaggccgagg ccgcctctgc ctctgagcta ttccagaagt 2640
agtgaggagg ctttttgga ggcctaggct tttgcaaaaa gctcccggga gcttgtatat 2700
ccattttcgg atctgatcag cacgtgttga caattaatca tcggcatagt atatcggcat 2760
agtataatac gacaaggtga ggaactaaac catggccaag cctttgtctc aagaagaatc 2820
caccctcatt gaaagagcaa cggctacaat caacagcatc cccatctctg aagactacag 2880
cgtcgccagc gcagctctct ctagcgacgg ccgcatcttc actggtgtca atgtatatca 2940
ttttactggg ggaccttgtg cagaactcgt ggtgctgggc actgctgctg ctgcggcagc 3000
tggcaacctg acttgtatcg tcgcgatcgg aaatgagaac aggggcatct tgagcccctg 3060
cggacggtgc cgacaggtgc ttctcgatct gcatcctggg atcaaagcca tagtgaagga 3120
cagtgatgga cagccgacgg cagttgggat tcgtgaattg ctgccctctg gttatgtgtg 3180
ggagggctaa gcacttcgtg gccgaggagc aggactgaca cgtgctacga gatttcgatt 3240
ccaccgccgc cttctatgaa aggttgggct tcggaatcgt tttccgggac gccggctgga 3300
tgatcctcca gcgcggggat ctcatgctgg agttcttcgc ccaccccaac ttgtttattg 3360
cagcttataa tggttacaaa taaagcaata gcatcacaaa tttcacaaat aaagcatttt 3420
tttcactgca ttctagttgt ggtttgtcca aactcatcaa tgtatcttat catgtctgta 3480
taccgtcgac ctctagctag agcttggcgt aatcatggtc atagctgttt cctgtgtgaa 3540
attgttatcc gctcacaatt ccacacaaca tacgagccgg aagcataaag tgtaaagcct 3600
ggggtgccta atgagtgagc taactcacat taattgcgtt gcgctcactg cccgctttcc 3660
agtcgggaaa cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg 3720
gtttgcgtat tgggcgctct tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc 3780

ggctgcggcg agcggtatca gctcactcaa aggcggtaat acggttatcc acagaatcag 3840
gggataacgc aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa 3900
aggccgcgtt gctggcgttt ttccataggc tccgcccccc tgacgagcat cacaaaaatc 3960
gacgctcaag tcagaggtgg cgaaacccga caggactata aagataccag gcgtttcccc 4020
ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga tacctgtccg 4080
cctttctccc ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt 4140
cggtgtaggt cgttcgctcc aagctgggct gtgtgcacga acccccgtt cagcccgacc 4200
gctgcgcctt atccggtaac tatcgtcttg agtccaaccc ggtaagacac gacttatcgc 4260
cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag 4320
agttcttgaa gtggtggcct aactacggct acactagaag aacagtattt ggtatctgcg 4380
ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa 4440
ccaccgctgg tagcggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat 4500
ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac 4560
gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt 4620
aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc 4680
aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg 4740
cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg 4800
ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc 4860
cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta 4920
ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg 4980
ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct 5040
ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta 5100
gctccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg 5160
ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga 5220
ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt 5280
gcccggcgtc aatacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca 5340
ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt 5400
cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt 5460
ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga 5520
aatgttgaat actcatactc ttccttttc aatattattg aagcatttat cagggttatt 5580
gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc 5640
gcacatttcc ccgaaaagtg ccacctgacg tc 5672

<210> 194

<211> 2850
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS8993

<400> 194
atggtggatc tacgcacgct cggctacagc cagcagcaac aggagaagat caaaccgaag 60
gttcgttcga cagtggcgca gcaccacgag gcactggtcg gccacgggtt tacacacgcg 120
cacatcgttg cgttaagcca acacccggca gcgttaggga ccgtcgctgt caagtatcag 180
gacatgatcg cagcgttgcc agaggcgaca cacgaagcga tcgttggcgt cggcaaacag 240
tggtccggcg cacgcgctct ggaggccttg ctcacggtgg cgggagagtt gagaggtcca 300
ccgttacagt tggacacagg ccaacttctc aagattgcaa aacgtggcgg cgtgaccgca 360
gtggaggcag tgcatgcatg gcgcaatgca ctgacgggtg ccccgctcaa cttgaccccg 420
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 480
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 540
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 600
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 660
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 720
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 780
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 840
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 900
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagcaatatt 960
ggtggcaagc aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac 1020
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg 1080
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag 1140
gtggtggcca tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg 1200
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 1260
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 1320
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 1380
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 1440

gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 1500
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 1560
gccagcaata ttggtggcaa gcaggcgctg gagacggtgc aggcgctgtt gccggtgctg 1620
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 1680
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 1740
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 1800
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 1860
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 1920
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 1980
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 2040
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg 2100
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac ccctcagcag 2160
gtggtggcca tcgccagcaa tggcggcggc aggccggcgc tggagagcat tgttgcccag 2220
ttatctcgcc ctgatccggg atccaagtct ggcatctacc agattaagaa caccctgaat 2280
aacaaagtgt atgttggatc tgccaaggac tttgagaaaa gatggaagag gcacttcaag 2340
gatctggaaa agggttgcca cagctccatc aagctccaaa ggtctttcaa caaacatggc 2400
aatgtctttg agtgctccat tttggaagag atcccttatg aaaaggacct cattatagag 2460
agggaaaact tctggatcaa ggagctgaac tcaaaaatca atggatacaa cattgcagat 2520
gccacatttg gtgacacctg ttctactcac ccccttaagg aagagatcat taaaaagagg 2580
tctgagactg tgaaggctaa gatgttgaaa ctggggcctg atggcagaaa ggcactctac 2640
tctaagccag gtagcaaaaa tggaaggtgg aaccctgaga cccacaagtt ctgcaaatgt 2700
ggggtgagaa tacagacatc agcctacact tgcagcaagt gtaggaacag atcaggagag 2760
aataactcct ttttcaacca caagcattct gacatcacca agtccaaaat ttcaggcaag 2820
atgaaaggtc aaggtccatc ggccgactga 2850

<210> 195

<211> 2811
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS8994

<400> 195
atggtggatc tacgcacgct cggctacagc cagcagcaac aggagaagat caaaccgaag 60
gttcgttcga cagtggcgca gcaccacgag gcactggtcg gccacgggtt tacacacgcg 120
cacatcgttg cgttaagcca acacccggca gcgttaggga ccgtcgctgt caagtatcag 180
gacatgatcg cagcgttgcc agaggcgaca cacgaagcga tcgttggcgt cggcaaacag 240
tggtccggcg cacgcgctct ggaggccttg ctcacggtgg cgggagagtt gagaggtcca 300
ccgttacagt tggacacagg ccaacttctc aagattgcaa aacgtggcgg cgtgaccgca 360
gtggaggcag tgcatgcatg gcgcaatgca ctgacgggtg ccccgctcaa cttgaccccg 420
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 480
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 540
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 600
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 660
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 720
acccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 780
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 840
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 900
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagcaatatt 960
ggtggcaagc aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac 1020
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg 1080
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag 1140
gtggtggcca tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg 1200
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 1260
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 1320
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 1380
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 1440
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 1500
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 1560
gccagcaata ttggtggcaa gcaggcgctg gagacggtgc aggcgctgtt gccggtgctg 1620
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 1680
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 1740
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 1800
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 1860
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 1920

```
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc   1980
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac   2040
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg   2100
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac ccctcagcag   2160
gtggtggcca tcgccagcaa tggcggcggc aggccggcgc tggagagcat tgttgcccag   2220
ttatctcgcc ctgatccggg atccaagtct ggcatctacc agattaagaa caccctgaat   2280
aacaaagtgt atgttggatc tgccaaggac tttgagaaaa gatggaagag gcacttcaag   2340
gatctggaaa agggttgcca cagctccatc aagctccaaa ggtctttcaa caaacatggc   2400
aatgtctttg agtgctccat tttggaagag atcccttatg aaaaggacct cattatagag   2460
agggaaaact tctggatcaa ggagctgaac tcaaaaatca atggatacaa cattgcagat   2520
gccacatttg gtgacacctg ttctactcac ccccttaagg aagagatcat taaaaagagg   2580
tctgagactg tgaaggctaa gatgttgaaa ctggggcctg atggcagaaa ggcactctac   2640
tctaagccag gtagcaaaaa tggaaggtgg aaccctgaga cccacaagtt ctgcaaatgt   2700
ggggtgagaa tacagacatc agcctacact tgcagcaagt gtaggaacag atcaggagag   2760
aataactcct ttttcaacca caagcattct caaggtccat cggccgactg a            2811


<210>  196

<211>  1321
<212>  PRT
<213>  Artificial sequence

<220>
<223>  pCLS7184

<400>  196
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1               5                   10                  15


Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
            20                  25                  30


Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
        35                  40                  45


Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
    50                  55                  60


Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65                  70                  75                  80


Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
                85                  90                  95


His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
            100                 105                 110


Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
        115                 120                 125


Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
    130                 135                 140


Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145                 150                 155                 160


Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
```

165 170 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
180 185 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
195 200 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
210 215 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225 230 235 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
245 250 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
260 265 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
275 280 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
290 295 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305 310 315 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
325 330 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
340 345 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
355 360 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
370 375 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385 390 395 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
405 410 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
420 425 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln

435 440 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
450 455 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465 470 475 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
485 490 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
500 505 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
515 520 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
530 535 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545 550 555 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
565 570 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
580 585 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
595 600 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
610 615 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625 630 635 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
645 650 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
660 665 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
675 680 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
690 695 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln

705 710 715 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
725 730 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
740 745 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
755 760 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
770 775 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785 790 795 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
805 810 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
820 825 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
835 840 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
850 855 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865 870 875 880

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
885 890 895

Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala
900 905 910

Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro
915 920 925

His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg
930 935 940

Thr Ser His Arg Val Ala Asp His Ala Gln Val Val Arg Val Leu Gly
945 950 955 960

Phe Phe Gln Cys His Ser His Pro Ala Gln Ala Phe Asp Asp Ala Met
965 970 975

Thr Gln Phe Gly Met Ser Arg His Gly Leu Leu Gln Leu Phe Arg Arg

980 985 990

Val Gly Val Thr Glu Leu Glu Ala Arg Ser Gly Thr Leu Pro Pro Ala
995 1000 1005

Ser Gln Arg Trp Asp Arg Ile Leu Gln Ala Ser Gly Met Lys Arg Ala
1010 1015 1020

Lys Pro Ser Pro Thr Ser Thr Gln Thr Pro Asp Gln Ala Ser Leu His
1025 1030 1035 1040

Ala Phe Ala Asp Ser Leu Glu Arg Asp Leu Asp Ala Pro Ser Pro Met
1045 1050 1055

His Glu Gly Asp Gln Thr Arg Ala Ser Ser Arg Lys Arg Ser Arg Ser
1060 1065 1070

Asp Arg Ala Val Thr Gly Pro Ser Ala Gln Gln Ser Phe Glu Val Arg
1075 1080 1085

Val Pro Glu Gln Arg Asp Ala Leu His Leu Pro Leu Ser Trp Arg Val
1090 1095 1100

Lys Arg Pro Arg Thr Ser Ile Gly Gly Gly Leu Pro Asp Pro Ile Ser
1105 1110 1115 1120

Arg Ser Gln Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu
1125 1130 1135

Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu
1140 1145 1150

Ile Ala Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met
1155 1160 1165

Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly
1170 1175 1180

Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp
1185 1190 1195 1200

Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu
1205 1210 1215

Pro Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn Gln
1220 1225 1230

Thr Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro
1235 1240 1245

Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys

1250 1255 1260

Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys
1265 1270 1275 1280

Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met
1285 1290 1295

Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn
1300 1305 1310

Asn Gly Glu Ile Asn Phe Ala Ala Asp
1315 1320

<210> 197

<211> 97
<212> PRT
<213> Artificial sequence

<220>
<223> I-TevI catalytic domain

<400> 197
Met Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val
1 5 10 15

Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe
20 25 30

Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser
35 40 45

Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile
50 55 60

Pro Tyr Ala Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys
65 70 75 80

Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe
85 90 95

Gly

<210> 198

<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> I-TevI deletion-intolerant region

<400> 198
Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys Arg

1 5 10 15

Ser

<210> 199

<211> 49
<212> PRT
<213> Artificial sequence

<220>
<223> I-TevI deletion-tolerant region

<400> 199
Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly Arg Lys
1 5 10 15

Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp Asn Pro Glu
20 25 30

Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr Ser Ala Tyr
35 40 45

Thr

<210> 200

<211> 7
<212> PRT
<213> Artificial sequence

<220>
<223> I-TevI zinc finger

<400> 200
Cys Ser Lys Cys Arg Asn Arg
1 5

<210> 201

<211> 8080
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS9021

<400> 201
aaaacgccag caacgcggcc tttttacggt tcctggcctt ttgctggcct tttgctcaca 60
tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagtgag 120
ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg 180
aagagcgcct gatgcggtat tttctcctta cgcatctgtg cggtatttca caccgcatat 240
atggtgcact ctcagtacaa tctgctctga tgccgcatag ttaagccagt atacactccg 300
ctatcgctac gtgactgggt catggctgcg ccccgacacc cgccaacacc cgctgacgcg 360
ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg 420
agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcgaggca gctgcggtaa 480
agctcatcag cgtggtcgtg aagcgattca cagatgtctg cctgttcatc cgcgtccagc 540
tcgttgagtt tctccagaag cgttaatgtc tggcttctga taaagcgggc catgttaagg 600

gcggtttttt cctgtttggt cactgatgcc tccgtgtaag ggggatttct gttcatgggg 660
gtaatgatac cgatgaaacg agagaggatg ctcacgatac gggttactga tgatgaacat 720
gcccggttac tggaacgttg tgagggtaaa caactggcgg tatggatgcg gcgggaccag 780
agaaaaatca ctcagggtca atgccagcgc ttcgttaata cagatgtagg tgttccacag 840
ggtagccagc agcatcctgc gatgcagatc cggaacataa tggtgcaggg cgctgacttc 900
cgcgtttcca gactttacga aacacggaaa ccgaagacca ttcatgttgt tgctcaggtc 960
gcagacgttt tgcagcagca gtcgcttcac gttcgctcgc gtatcggtga ttcattctgc 1020
taaccagtaa ggcaaccccg ccagcctagc cgggtcctca acgacaggag cacgatcatg 1080
cgcacccgtg gggccgccat gccggcgata atggcctgct tctcgccgaa acgtttggtg 1140
gcgggaccag tgacgaaggc ttgagcgagg gcgtgcaaga ttccgaatac cgcaagcgac 1200
aggccgatca tcgtcgcgct ccagcgaaag cggtcctcgc cgaaaatgac ccagagcgct 1260
gccggcacct gtcctacgag ttgcatgata aagaagacag tcataagtgc ggcgacgata 1320
gtcatgcccc gcgcccaccg gaaggagctg actgggttga aggctctcaa gggcatcggt 1380
cgagatcccg gtgcctaatg agtgagctaa cttacattaa ttgcgttgcg ctcactgccc 1440
gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg 1500
agaggcggtt tgcgtattgg gcgccagggt ggtttttctt ttcaccagtg agacgggcaa 1560
cagctgattg cccttcaccg cctggccctg agagagttgc agcaagcggt ccacgctggt 1620
ttgccccagc aggcgaaaat cctgtttgat ggtggttaac ggcgggatat aacatgagct 1680
gtcttcggta tcgtcgtatc ccactaccga gatatccgca ccaacgcgca gcccggactc 1740
ggtaatggcg cgcattgcgc ccagcgccat ctgatcgttg gcaaccagca tcgcagtggg 1800
aacgatgccc tcattcagca tttgcatggt ttgttgaaaa ccggacatgg cactccagtc 1860
gccttcccgt tccgctatcg gctgaatttg attgcgagtg agatatttat gccagccagc 1920
cagacgcaga cgcgccgaga cagaacttaa tgggcccgct aacagcgcga tttgctggtg 1980
acccaatgcg accagatgct ccacgcccag tcgcgtaccg tcttcatggg agaaaataat 2040
actgttgatg ggtgtctggt cagagacatc aagaaataac gccggaacat tagtgcaggc 2100
agcttccaca gcaatggcat cctggtcatc cagcggatag ttaatgatca gcccactgac 2160
gcgttgcgcg agaagattgt gcaccgccgc tttacaggct tcgacgccgc ttcgttctac 2220
catcgacacc accacgctgg cacccagttg atcggcgcga gatttaatcg ccgcgacaat 2280
ttgcgacggc gcgtgcaggg ccagactgga ggtggcaacg ccaatcagca acgactgttt 2340
gcccgccagt tgttgtgcca cgcggttggg aatgtaattc agctccgcca tcgccgcttc 2400
cacttttcc cgcgttttcg cagaaacgtg gctggcctgg ttcaccacgc gggaaacggt 2460
ctgataagag acaccggcat actctgcgac atcgtataac gttactggtt tcacattcac 2520
caccctgaat tgactctctt ccgggcgcta tcatgccata ccgcgaaagg ttttgcgcca 2580
ttcgatggtg tccgggatct cgacgctctc ccttatgcga ctcctgcatt aggaagcagc 2640
ccagtagtag gttgaggccg ttgagcaccg ccgccgcaag gaatggtgca tgcaaggaga 2700
tggcgcccaa cagtcccccg gccacggggc ctgccaccat acccacgccg aaacaagcgc 2760
tcatgagccc gaagtggcga gcccgatctt ccccatcggt gatgtcggcg atataggcgc 2820
cagcaaccgc acctgtggcg ccggtgatgc cggccacgat gcgtccggcg tagaggatcg 2880
agatctcgat cccgcgaaat taatacgact cactataggg gaattgtgag cggataacaa 2940
ttcccctcta gaaataattt tgtttaactt taagaaggag atataccatg gtggatctac 3000
gcacgctcgg ctacagccag cagcaacagg agaagatcaa accgaaggtt cgttcgacag 3060
tggcgcagca ccacgaggca ctggtcggcc acgggtttac acacgcgcac atcgttgcgt 3120
taagccaaca cccggcagcg ttagggaccg tcgctgtcaa gtatcaggac atgatcgcag 3180
cgttgccaga ggcgacacac gaagcgatcg ttggcgtcgg caaacagtgg tccggcgcac 3240
gcgctctgga ggccttgctc acggtggcgg gagagttgag aggtccaccg ttacagttgg 3300
acacaggcca acttctcaag attgcaaaac gtggcggcgt gaccgcagtg gaggcagtgc 3360
atgcatggcg caatgcactg acgggtgccc cgctcaactt gaccccggag caggtggtgg 3420
ccatcgccag ccacgatggc ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg 3480
tgctgtgcca ggcccacggc ttgacccccc agcaggtggt ggccatcgcc agcaatggcg 3540
gtggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc caggcccacg 3600
gcttgacccc ggagcaggtg gtggccatcg ccagcaatat tggtggcaag caggcgctgg 3660
agacggtgca ggcgctgttg ccggtgctgt gccaggccca cggcttgacc ccccagcagg 3720
tggtggccat cgccagcaat ggcggtggca agcaggcgct ggagacggtc cagcggctgt 3780
tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc atcgccagca 3840
atattggtgg caagcaggcg ctggagacgg tgcaggcgct gttgccggtg ctgtgccagg 3900
cccacggctt gaccccggag caggtggtgg ccatcgccag caatattggt ggcaagcagg 3960
cgctggagac ggtgcaggcg ctgttgccgg tgctgtgcca ggcccacggc ttgaccccgg 4020
agcaggtggt ggccatcgcc agcaatattg gtggcaagca ggcgctggag acggtgcagg 4080
cgctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg gtggccatcg 4140
ccagccacga tggcggcaag caggcgctgg agacggtcca gcggctgttg ccggtgctgt 4200
gccaggccca cggcttgacc ccggagcagg tggtggccat cgccagccac gatggcggca 4260
agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc cacggcttga 4320
ccccccagca ggtggtggcc atcgccagca atggcggtgg caagcaggcg ctggagacgg 4380
tccagcggct gttgccggtg ctgtgccagg cccacggctt gaccccggag caggtggtgg 4440
ccatcgccag caatattggt ggcaagcagg cgctggagac ggtgcaggcg ctgttgccgg 4500
tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc agcaatattg 4560
gtggcaagca ggcgctggag acggtgcagg cgctgttgcc ggtgctgtgc caggcccacg 4620
gcttgacccc ggagcaggtg gtggccatcg ccagccacga tggcggcaag caggcgctgg 4680

```
agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc ccggagcagg   4740
tggtggccat cgccagccac gatggcggca agcaggcgct ggagacggtc cagcggctgt   4800
tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc atcgccagcc   4860
acgatggcgg caagcaggcg ctggagacgg tccagcggct gttgccggtg ctgtgccagg   4920
cccacggctt gacccccag caggtggtgg ccatcgccag caatggcggt ggcaagcagg   4980
cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc ttgaccccgg   5040
agcaggtggt ggccatcgcc agccacgatg gcggcaagca ggcgctggag acggtccagc   5100
ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc tcagcaggtg gtggccatcg   5160
ccagcaatgg cggcggcagg ccggcgctgg agagcattgt tgcccagtta tctcgccctg   5220
atccgggatc caagtctggc atctaccaga ttaagaacac cctgaataac aaagtgtatg   5280
ttggatctgc caaggacttt gagaaaagat ggaagaggca cttcaaggat ctggaaaagg   5340
gttgccacag ctccatcaag ctccaaaggt ctttcaacaa acatggcaat gtctttgagt   5400
gctccatttt ggaagagatc ccttatgaaa aggacctcat tatagagagg gaaaacttct   5460
ggatcaagga gctgaactca aaaatcaatg gatacaacat tgcagatgcc acatttggtg   5520
acacctgttc tactcacccc cttaaggaag agatcattaa aaagaggtct gagactgtga   5580
aggctaagat gttgaaactg gggcctgatg gcagaaaggc actctactct aagccaggta   5640
gcaaaaatgg aaggtggaac cctgagaccc acaagttctg caaatgtggg gtgagaatac   5700
agacatcagc ctacacttgc agcaagtgta ggaacagatc aggagagaat aactcctttt   5760
tcaaccacaa gcattctgac atcaccaagt ccaaaatttc aggcaagatg aaaggtcaag   5820
gtccatcggc cgcactcgag caccaccacc accaccactg agatccggct gctaacaaag   5880
cccgaaagga agctgagttg gctgctgcca ccgctgagca ataactagca taacccccttg   5940
gggcctctaa acgggtcttg aggggttttt tgctgaaagg aggaactata tccggattgg   6000
cgaatgggac gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg ttacgcgcag   6060
cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct tcccttcctt   6120
tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc ctttagggtt   6180
ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg atggttcacg   6240
tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttggagt ccacgttctt   6300
taatagtgga ctcttgttcc aaactggaac aacactcaac cctatctcgg tctattcttt   6360
tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc tgatttaaca   6420
aaaatttaac gcgaatttta acaaaatatt aacgtttaca atttcaggtg gcacttttcg   6480
gggaaatgtg cgcggaaccc ctatttgttt atttttctaa atacattcaa atatgtatcc   6540
gctcatgaat taattcttag aaaaactcat cgagcatcaa atgaaactgc aatttattca   6600
tatcaggatt atcaatacca tatttttgaa aaagccgttt ctgtaatgaa ggagaaaact   6660
caccgaggca gttccatagg atggcaagat cctggtatcg gtctgcgatt ccgactcgtc   6720
caacatcaat acaacctatt aatttccct cgtcaaaaat aaggttatca agtgagaaat   6780
caccatgagt gacgactgaa tccggtgaga atggcaaaag tttatgcatt tctttccaga   6840
cttgttcaac aggccagcca ttacgctcgt catcaaaatc actcgcatca accaaaccgt   6900
tattcattcg tgattgcgcc tgagcgagac gaaatacgcg atcgctgtta aaaggacaat   6960
tacaaacagg aatcgaatgc aaccggcgca ggaacactgc cagcgcatca acaatatttt   7020
cacctgaatc aggatattct tctaatacct ggaatgctgt tttcccgggg atcgcagtgg   7080
tgagtaacca tgcatcatca ggagtacgga taaaatgctt gatggtcgga agaggcataa   7140
attccgtcag ccagtttagt ctgaccatct catctgtaac atcattggca acgctacctt   7200
tgccatgttt cagaaacaac tctggcgcat cgggcttccc atacaatcga tagattgtcg   7260
cacctgattg cccgacatta tcgcgagccc atttataccc atataaatca gcatccatgt   7320
tggaatttaa tcgcggccta gagcaagacg tttcccgttg aatatggctc ataacacccc   7380
ttgtattact gtttatgtaa gcagacagtt ttattgttca tgaccaaaat cccttaacgt   7440
gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat   7500
ccttttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg   7560
gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga   7620
gcgcagatac caaatactgt ccttctagtg tagccgtagt taggccacca cttcaagaac   7680
tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt   7740
ggcgataagt cgtgtcttac cgggttggac tcaagacgat agttaccgga taaggcgcag   7800
cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc   7860
gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaag   7920
gcggacaggt atccggtaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca   7980
gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt   8040
cgatttttgt gatgctcgtc agggggggcgg agcctatgga                        8080
```

<210> 202

<211> 8041
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS9022

<400> 202

aaaacgccag caacgcggcc tttttacggt tcctggcctt ttgctggcct tttgctcaca 60
tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagtgag 120
ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg 180
aagagcgcct gatgcggtat tttctcctta cgcatctgtg cggtatttca caccgcatat 240
atggtgcact ctcagtacaa tctgctctga tgccgcatag ttaagccagt atacactccg 300
ctatcgctac gtgactgggt catggctgcg ccccgacacc cgccaacacc cgctgacgcg 360
ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg 420
agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcgaggca gctgcggtaa 480
agctcatcag cgtggtcgtg aagcgattca cagatgtctg cctgttcatc cgcgtccagc 540
tcgttgagtt tctccagaag cgttaatgtc tggcttctga taaagcgggc catgttaagg 600
gcggtttttt cctgtttggt cactgatgcc tccgtgtaag ggggatttct gttcatgggg 660
gtaatgatac cgatgaaacg agagaggatg ctcacgatac gggttactga tgatgaacat 720
gcccggttac tggaacgttg tgagggtaaa caactggcgg tatggatgcg gcgggaccag 780
agaaaaatca ctcagggtca atgccagcgc ttcgttaata cagatgtagg tgttccacag 840
ggtagccagc agcatcctgc gatgcagatc cggaacataa tggtgcaggg cgctgacttc 900
cgcgtttcca gactttacga aacacggaaa ccgaagacca ttcatgttgt tgctcaggtc 960
gcagacgttt tgcagcagca gtcgcttcac gttcgctcgc gtatcggtga ttcattctgc 1020
taaccagtaa ggcaaccccg ccagcctagc cgggtcctca acgacaggag cacgatcatg 1080
cgcacccgtg gggccgccat gccggcgata atggcctgct tctcgccgaa acgtttggtg 1140
gcgggaccag tgacgaaggc ttgagcgagg gcgtgcaaga ttccgaatac cgcaagcgac 1200
aggccgatca tcgtcgcgct ccagcgaaag cggtcctcgc cgaaaatgac ccagagcgct 1260
gccggcacct gtcctacgag ttgcatgata aagaagacag tcataagtgc ggcgacgata 1320
gtcatgcccc gcgcccaccg gaaggagctg actgggttga aggctctcaa gggcatcggt 1380
cgagatcccg gtgcctaatg agtgagctaa cttacattaa ttgcgttgcg ctcactgccc 1440
gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg 1500
agaggcggtt tgcgtattgg gcgccagggt ggtttttctt ttcaccagtg agacgggcaa 1560
cagctgattg cccttcaccg cctggccctg agagagttgc agcaagcggt ccacgctggt 1620
ttgccccagc aggcgaaaat cctgtttgat ggtggttaac ggcgggatat aacatgagct 1680
gtcttcggta tcgtcgtatc ccactaccga gatatccgca ccaacgcgca gcccggactc 1740
ggtaatggcg cgcattgcgc ccagcgccat ctgatcgttg gcaaccagca tcgcagtggg 1800
aacgatgccc tcattcagca tttgcatggt ttgttgaaaa ccggacatgg cactccagtc 1860
gccttcccgt tccgctatcg gctgaatttg attgcgagtg agatatttat gccagccagc 1920
cagacgcaga cgcgccgaga cagaacttaa tgggcccgct aacagcgcga tttgctggtg 1980
acccaatgcg accagatgct ccacgcccag tcgcgtaccg tcttcatggg agaaaataat 2040
actgttgatg ggtgtctggt cagagacatc aagaaataac gccggaacat tagtgcaggc 2100
agcttccaca gcaatggcat cctggtcatc cagcggatag ttaatgatca gcccactgac 2160
gcgttgcgcg agaagattgt gcaccgccgc tttacaggct tcgacgccgc ttcgttctac 2220
catcgacacc accacgctgg cacccagttg atcggcgcga gatttaatcg ccgcgacaat 2280
ttgcgacggc gcgtgcaggg ccagactgga ggtggcaacg ccaatcagca acgactgttt 2340
gcccgccagt tgttgtgcca cgcggttggg aatgtaattc agctccgcca tcgccgcttc 2400
cacttttcc cgcgttttcg cagaaacgtg gctggcctgg ttcaccacgc gggaaacggt 2460
ctgataagag acaccggcat actctgcgac atcgtataac gttactggtt tcacattcac 2520
caccctgaat tgactctctt ccgggcgcta tcatgccata ccgcgaaagg ttttgcgcca 2580
ttcgatggtg tccgggatct cgacgctctc ccttatgcga ctcctgcatt aggaagcagc 2640
ccagtagtag gttgaggccg ttgagcaccg ccgccgcaag gaatggtgca tgcaaggaga 2700
tggcgcccaa cagtcccccg gccacggggc ctgccaccat acccacgccg aaacaagcgc 2760
tcatgagccc gaagtggcga gcccgatctt ccccatcggt gatgtcggcg atataggcgc 2820
cagcaaccgc acctgtggcg ccggtgatgc cggccacgat gcgtccggcg tagaggatcg 2880
agatctcgat cccgcgaaat taatacgact cactataggg gaattgtgag cggataacaa 2940
ttcccctcta gaaataattt tgtttaactt taagaaggag atataccatg gtggatctac 3000
gcacgctcgg ctacagccag cagcaacagg agaagatcaa accgaaggtt cgttcgacag 3060
tggcgcagca ccacgaggca ctggtcggcc acgggtttac acacgcgcac atcgttgcgt 3120
taagccaaca cccggcagcg ttagggaccg tcgctgtcaa gtatcaggac atgatcgcag 3180
cgttgccaga ggcgacacac gaagcgatcg ttggcgtcgg caaacagtgg tccggcgcac 3240
gcgctctgga ggccttgctc acggtggcgg gagagttgag aggtccaccg ttacagttgg 3300
acacaggcca acttctcaag attgcaaaac gtggcggcgt gaccgcagtg gaggcagtgc 3360
atgcatggcg caatgcactg acgggtgccc cgctcaactt gaccccggag caggtggtgg 3420
ccatcgccag ccacgatggc ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg 3480
tgctgtgcca ggcccacggc ttgacccccc agcaggtggt ggccatcgcc agcaatggcg 3540
gtggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc caggcccacg 3600
gcttgacccc ggagcaggtg gtggccatcg ccagcaatat tggtggcaag caggcgctgg 3660
agacggtgca ggcgctgttg ccggtgctgt gccaggccca cggcttgacc ccccagcagg 3720
tggtggccat cgccagcaat ggcggtggca agcaggcgct ggagacggtc cagcggctgt 3780
tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc atcgccagca 3840
atattggtgg caagcaggcg ctggagacgg tgcaggcgct gttgccggtg ctgtgccagg 3900
cccacggctt gaccccggag caggtggtgg ccatcgccag caatattggt ggcaagcagg 3960
cgctggagac ggtgcaggcg ctgttgccgg tgctgtgcca ggcccacggc ttgaccccgg 4020

agcaggtggt ggccatcgcc agcaatattg gtggcaagca ggcgctggag acggtgcagg 4080
cgctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg gtggccatcg 4140
ccagccacga tggcggcaag caggcgctgg agacggtcca gcggctgttg ccggtgctgt 4200
gccaggccca cggcttgacc ccggagcagg tggtggccat cgccagccac gatggcggca 4260
agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc cacggcttga 4320
ccccccagca ggtggtggcc atcgccagca atggcggtgg caagcaggcg ctggagacgg 4380
tccagcggct gttgccggtg ctgtgccagg cccacggctt gaccccggag caggtggtgg 4440
ccatcgccag caatattggt ggcaagcagg cgctggagac ggtgcaggcg ctgttgccgg 4500
tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc agcaatattg 4560
gtggcaagca ggcgctggag acggtgcagg cgctgttgcc ggtgctgtgc caggcccacg 4620
gcttgacccc ggagcaggtg gtggccatcg ccagccacga tggcggcaag caggcgctgg 4680
agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc ccggagcagg 4740
tggtggccat cgccagccac gatggcggca agcaggcgct ggagacggtc cagcggctgt 4800
tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc atcgccagcc 4860
acgatggcgg caagcaggcg ctggagacgg tccagcggct gttgccggtg ctgtgccagg 4920
cccacggctt gacccccag caggtggtgg ccatcgccag caatggcggt ggcaagcagg 4980
cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc ttgaccccgg 5040
agcaggtggt ggccatcgcc agccacgatg gcggcaagca ggcgctggag acggtccagc 5100
ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc tcagcaggtg gtggccatcg 5160
ccagcaatgg cggcggcagg ccggcgctgg agagcattgt tgcccagtta tctcgccctg 5220
atccgggatc caagtctggc atctaccaga ttaagaacac cctgaataac aaagtgtatg 5280
ttggatctgc caaggacttt gagaaaagat ggaagaggca cttcaaggat ctggaaaagg 5340
gttgccacag ctccatcaag ctccaaaggt ctttcaacaa acatggcaat gtctttgagt 5400
gctccatttt ggaagagatc ccttatgaaa aggacctcat tatagagagg gaaaacttct 5460
ggatcaagga gctgaactca aaaatcaatg gatacaacat tgcagatgcc acatttggtg 5520
acacctgttc tactcacccc cttaaggaag agatcattaa aaagaggtct gagactgtga 5580
aggctaagat gttgaaactg gggcctgatg gcagaaaggc actctactct aagccaggta 5640
gcaaaaatgg aaggtggaac cctgagaccc acaagttctg caaatgtggg gtgagaatac 5700
agacatcagc ctacacttgc agcaagtgta ggaacagatc aggagagaat aactcctttt 5760
tcaaccacaa gcattctcaa ggtccatcgg ccgcactcga gcaccaccac caccaccact 5820
gagatccggc tgctaacaaa gcccgaaagg aagctgagtt ggctgctgcc accgctgagc 5880
aataactagc ataaccccct ggggcctcta aacgggtctt gaggggtttt ttgctgaaag 5940
gaggaactat atccggattg gcgaatggga cgcgccctgt agcggcgcat taagcgcggc 6000
gggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc agcgccctag cgcccgctcc 6060
tttcgctttc ttcccttcct ttctcgccac gttcgccggc tttccccgtc aagctctaaa 6120
tcgggggctc cctttagggt tccgatttag tgctttacgg cacctcgacc ccaaaaaact 6180
tgattagggt gatggttcac gtagtgggcc atcgccctga tagacggttt ttcgcccttt 6240
gacgttggag tccacgttct ttaatagtgg actcttgttc caaactggaa caacactcaa 6300
ccctatctcg gtctattctt ttgatttata agggattttg ccgatttcgg cctattggtt 6360
aaaaaatgag ctgatttaac aaaaatttaa cgcgaatttt aacaaaatat taacgtttac 6420
aatttcaggt ggcacttttc ggggaaatgt gcgcggaacc cctatttgtt tatttttcta 6480
aatacattca aatatgtatc cgctcatgaa ttaattctta gaaaaactca tcgagcatca 6540
aatgaaactg caatttattc atatcaggat tatcaatacc atatttttga aaaagccgtt 6600
tctgtaatga aggagaaaac tcaccgaggc agttccatag gatggcaaga tcctggtatc 6660
ggtctgcgat tccgactcgt ccaacatcaa tacaacctat taatttcccc tcgtcaaaaa 6720
taaggttatc aagtgagaaa tcaccatgag tgacgactga atccggtgag aatggcaaaa 6780
gtttatgcat ttctttccag acttgttcaa caggccagcc attacgctcg tcatcaaaat 6840
cactcgcatc aaccaaaccg ttattcattc gtgattgcgc ctgagcgaga cgaaatacgc 6900
gatcgctgtt aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc aggaacactg 6960
ccagcgcatc aacaatattt tcacctgaat caggatattc ttctaatacc tggaatgctg 7020
ttttcccggg gatcgcagtg gtgagtaacc atgcatcatc aggagtacgg ataaaatgct 7080
tgatggtcgg aagaggcata aattccgtca gccagtttag tctgaccatc tcatctgtaa 7140
catcattggc aacgctacct ttgccatgtt tcagaaacaa ctctggcgca tcgggcttcc 7200
catacaatcg atagattgtc gcacctgatt gcccgacatt atcgcgagcc catttatacc 7260
catataaatc agcatccatg ttggaattta atcgcggcct agagcaagac gtttcccgtt 7320
gaatatggct cataacaccc cttgtattac tgtttatgta agcagacagt tttattgttc 7380
atgaccaaaa tcccttaacg tgagttttcg ttccactgag cgtcagaccc cgtagaaaag 7440
atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa 7500
aaaccaccgc taccagcggt ggtttgtttg ccggatcaag agctaccaac tctttttccg 7560
aaggtaactg gcttcagcag agcgcagata ccaaatactg tccttctagt gtagccgtag 7620
ttaggccacc acttcaagaa ctctgtagca ccgcctacat acctcgctct gctaatcctg 7680
ttaccagtgg ctgctgccag tggcgataag tcgtgtctta ccgggttgga ctcaagacga 7740
tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc 7800
ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagctatg agaaagcgcc 7860
acgcttcccg aagggagaaa ggcggacagg tatccggtaa gcggcagggt cggaacagga 7920
gagcgcacga gggagcttcc agggggaaac gcctggtatc tttatagtcc tgtcgggttt 7980
cgccacctct gacttgagcg tcgatttttg tgatgctcgt caggggggcg gagcctatgg 8040
a 8041

<210> 203

<211> 9363
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS6615

<400> 203
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 60
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 120
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 180
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 240
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 300
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 360
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 420
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 480
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 540
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 600
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 660
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 720
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 780
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 840
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 900
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 960
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1020
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1080
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1140
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1200
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1260
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1320
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1380
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1440
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1500
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1560
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1620
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1680
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1740
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1800
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1860
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1920
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 1980
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2040
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2100
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2160
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2220
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2280
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2340
gacacgcaaa cacaaataca cagcggcctt gccaccatgg ccaagtctgg catctaccag 2400
attaagaaca ccctgaataa caaagtgtat gttggatctg ccaaggactt tgagaaaaga 2460
tggaagaggc acttcaagga tctggaaaag ggttgccaca gctccatcaa gctccaaagg 2520
tctttcaaca aacatggcaa tgtctttgag tgctccattt tggaagagat cccttatgaa 2580
aaggacctca ttatagagag ggaaaacttc tggatcaagg agctgaactc aaaaatcaat 2640
ggatacaaca ttgcagatgc cacatttggt gacacctgtt ctactcaccc ccttaaggaa 2700
gagatcatta aaaagaggtc tgagactgtg aaggctaaga tgttgaaact ggggcctgat 2760
ggcagaaagg cactctactc taagccaggt agcaaaaatg gaaggtggaa ccctgagacc 2820
cacaagttct gcaaatgtgg ggtgagaata cagacatcag cctacacttg cagcaagtgt 2880
aggaacagat caggagagaa taactccttt ttcaaccaca agcattctca aggtccatct 2940
ggcaatacca aatataacaa agagttcctg ctgtacctgg ccggctttgt ggacggtaac 3000
ggtagcatca tcgctcagat taaaccaaac cagtcttata agtttaaaca tcagctaagc 3060
ttgacctttc aggtgactca aaagacccag cgccgttggt ttctggacaa actagtggat 3120
gaaattggcg ttggttacgt acgtgatcgc ggatccgttt ccgattacat tttaagcgaa 3180
atcaagccgc tgcacaactt cctgactcaa ctgcagccgt ttctgaaact gaaacagaaa 3240
caggcaaacc tggttctgaa aattatcgaa cagctgccgt ctgcaaaaga atccccggac 3300
aaattcctgg aagtttgtac ctgggtggat cagattgcag ctctgaacga ttctaagacg 3360

cgtaaaacca cttctgaaac cgttcgtgct gtgctggaca gcctgagcga gaagaagaaa 3420
tcctccccgg cggccgactg ataactcgag cgatcctcta gacgagctcc tcgagcctgc 3480
agcagctgaa gctttggact tcttcgccag aggtttggtc aagtctccaa tcaaggttgt 3540
cggcttgtct accttgccag aaatttacga aaagatggaa aagggtcaaa tcgttggtag 3600
atacgttgtt gacacttcta ataagcgaa tttcttatga tttatgattt ttattattaa 3660
ataagttata aaaaaaataa gtgtatacaa attttaaagt gactcttagg ttttaaaacg 3720
aaaattctta ttcttgagta actctttcct gtaggtcagg ttgctttctc aggtatagca 3780
tgaggtcgct cttattgacc acacctctac cggcatgcaa gcttggcgta atcatggtca 3840
tagctgtttc ctgtgtgaaa ttgttatccg ctcacaattc cacacaacat acgagccgga 3900
agcataaagt gtaaagcctg gggtgcctaa tgagtgagct aactcacatt aattgcgttg 3960
cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agcagatcta ttacattatg 4020
ggtggtatgt tggaataaaa atcaactatc atctactaac tagtatttac gttactagta 4080
tattatcata tacggtgtta gaagatgacg caaatgatga gaaatagtca tctaaattag 4140
tggaagctga aacgcaagga ttgataatgt aataggatca atgaatatta acatataaaa 4200
tgatgataat aatatttata gaattgtgta gaattgcaga ttccctttta tggattccta 4260
aatcctcgag gagaacttct agtatatcta catacctaat attattgcct tattaaaaat 4320
ggaatcccaa caattacatc aaaatccaca ttctcttcaa aatcaattgt cctgtacttc 4380
cttgttcatg tgtgttcaaa aacgttatat ttataggata attatactct atttctcaac 4440
aagtaattgg ttgtttggcc gagcggtcta aggcgcctga ttcaagaaat atcttgaccg 4500
cagttaactg tgggaatact caggtatcgt aagatgcaag agttcgaatc tcttagcaac 4560
cattattttt ttcctcaaca taacgagaac acacaggggc gctatcgcac agaatcaaat 4620
tcgatgactg gaaatttttt gttaatttca gaggtcgcct gacgcatata ccttttcaa 4680
ctgaaaaatt gggagaaaaa ggaaaggtga gagccgcgga accggctttt catatagaat 4740
agagaagcgt tcatgactaa atgcttgcat cacaatactt gaagttgaca atattattta 4800
aggacctatt gtttttcca ataggtggtt agcaatcgtc ttactttcta acttttctta 4860
ccttttacat ttcagcaata tatatatata tatttcaagg atataccatt ctaatgtctg 4920
cccctaagaa gatcgtcgtt ttgccaggtg accacgttgg tcaagaaatc acagccgaag 4980
ccattaaggt tcttaaagct atttctgatg ttcgttccaa tgtcaagttc gatttcgaaa 5040
atcatttaat tggtggtgct gctatcgatg ctacaggtgt cccacttcca gatgaggcgc 5100
tggaagcctc caagaaggtt gatgccgttt tgttaggtgc tgtgggtggt cctaaatggg 5160
gtaccggtag tgttagacct gaacaaggtt tactaaaaat ccgtaaagaa cttcaattgt 5220
acgccaactt aagaccatgt aactttgcat ccgactctct tttagactta tctccaatca 5280
agccacaatt tgctaaaggt actgacttcg ttgttgtcag agaattagtg ggaggtattt 5340
actttggtaa gagaaaggaa gacgatggtg atggtgtcgc ttgggatagt gaacaataca 5400
ccgttccaga agtgcaaaga atcacaagaa tggccgcttt catggcccta caacatgagc 5460
caccattgcc tatttggtcc ttggataaag ctaatgtttt ggcctcttca agattatgga 5520
gaaaaactgt ggaggaaacc atcaagaacg aattccctac attgaaggtt caacatcaat 5580
tgattgattc tgccgccatg atcctagtta agaacccaac ccacctaaat ggtattataa 5640
tcaccagcaa catgtttggt gatatcatct ccgatgaagc ctccgttatc ccaggttcct 5700
tgggtttgtt gccatctgcg tccttggcct ctttgccaga caagaacacc gcatttggtt 5760
tgtacgaacc atgccacggt tctgctccag atttgccaaa gaataaggtc aaccctatcg 5820
ccactatctt gtctgctgca atgatgttga aattgtcatt gaacttgcct gaagaaggta 5880
aggccattga agatgcagtt aaaaaggttt tggatgcagg tatcagaact ggtgatttag 5940
gtggttccaa cagtaccacg gaagtcggtg atgctgtcgc cgaagaagtt aagaaaatcc 6000
ttgcttaaaa agattctctt tttttatgat atttgtacat aaactttata aatgaaattc 6060
ataatagaaa cgacacgaaa ttacaaaatg gaatatgttc atagggtaga cgaaactata 6120
tacgcaatct acatacattt atcaagaagg agaaaaagga ggatgtaaag gaatacaggt 6180
aagcaaattg atactaatgg ctcaacgtga taaggaaaaa gaattgcact ttaacattaa 6240
tattgacaag gaggagggca ccacacaaaa agttaggtgt aacagaaaat catgaaacta 6300
tgattcctaa tttatatatt ggaggatttt ctctaaaaaa aaaaaaatac aacaaataaa 6360
aaacactcaa tgacctgacc atttgatgga gtttaagtca ataccttctt gaaccatttc 6420
ccataatggt gaaagttccc tcaagaattt tactctgtca gaaacggcct taacgacgta 6480
gtcgacctcc tcttcagtac taaatctacc aataccaaat ctgatggaag aatgggctaa 6540
tgcatcatcc ttacccagcg catgtaaaac ataagaaggt tctagggaag cagatgtaca 6600
ggctgaaccc gaggataatg cgatatccct tagtgccatc aataaagatt ctccttccac 6660
gtaggcgaaa gaaacgttaa cacaccctgg ataacgatga tctggagatc cgttcaacgt 6720
ggtatgttca gcggataata gacctttgac taatttatcg gatagtcttt tgatgtgagc 6780
ttggtcgttg tcaaattctt tcttcatcaa tctcgcagct tcaccaaatc ccgctaccaa 6840
tgggggggcc aaagtaccag atctgctgca ttaatgaatc ggccaacgcg cggggagagg 6900
cggtttgcgt attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt 6960
tcggctgcgg cgagcggtat cagcatcgat gaattccacg gactatagac tatactagta 7020
tactccgtct actgtacgat acacttccgc tcaggtcctt gtcctttaac gaggccttac 7080
cactcttttg ttactctatt gatccagctc agcaaaggca gtgtgatcta agattctatc 7140
ttcgcgatgt agtaaaacta gctagaccga gaaagagact agaaatgcaa aaggcacttc 7200
tacaatggct gccatcatta ttatccgatg tgacgctgca gcttctcaat gatattcgaa 7260
tacgctttga ggagatacag cctaatatcc gacaaactgt tttacagatt tacgatcgta 7320
cttgttaccc atcattgaat tttgaacatc cgaacctggg agttttccct gaaacagata 7380
gtatatttga acctgtataa taatatatag tctagcgctt tacggaagac aatgtatgta 7440

tttcggttcc tggagaaact attgcatcta ttgcataggt aatcttgcac gtcgcatccc 7500
cggttcattt tctgcgtttc catcttgcac ttcaatagca tatctttgtt aacgaagcat 7560
ctgtgcttca ttttgtagaa caaaaatgca acgcgagagc gctaattttt caaacaaaga 7620
atctgagctg catttttaca gaacagaaat gcaacgcgaa agcgctattt taccaacgaa 7680
gaatctgtgc ttcatttttg taaaacaaaa atgcaacgcg agagcgctaa tttttcaaac 7740
aaagaatctg agctgcattt ttacagaaca gaaatgcaac gcgagagcgc tattttacca 7800
acaaagaatc tatacttctt ttttgttcta caaaaatgca tcccgagagc gctatttttc 7860
taacaaagca tcttagatta ctttttttct cctttgtgcg ctctataatg cagtctcttg 7920
ataacttttt gcactgtagg tccgttaagg ttagaagaag gctactttgg tgtctatttt 7980
ctcttccata aaaaaagcct gactccactt cccgcgttta ctgattacta gcgaagctgc 8040
gggtgcattt tttcaagata aaggcatccc cgattatatt ctataccgat gtggattgcg 8100
catactttgt gaacagaaag tgatagcgtt gatgattctt cattggtcag aaaattatga 8160
acggtttctt ctattttgtc tctatatact acgtatagga aatgtttaca ttttcgtatt 8220
gttttcgatt cactctatga atagttctta ctacaatttt tttgtctaaa gagtaatact 8280
agagataaac ataaaaaatg tagaggtcga gtttagatgc aagttcaagg agcgaaaggt 8340
ggatgggtag gttatatagg gatatagcac agagatatat agcaaagaga tacttttgag 8400
caatgtttgt ggaagcggta ttcgcaatat tttagtagct cgttacagtc cggtgcgttt 8460
ttggtttttt gaaagtgcgt cttcagagcg cttttggttt tcaaaagcgc tctgaagttc 8520
ctatactttc tagagaatag gaacttcgga ataggaactt caaagcgttt ccgaaaacga 8580
gcgcttccga aaatgcaacg cgagctgcgc acatacagct cactgttcac gtcgcaccta 8640
tatctgcgtg ttgcctgtat atatatatac atgagaagaa cggcatagtg cgtgtttatg 8700
cttaaatgcg tacttatatg cgtctattta tgtaggatga aaggtagtct agtacctcct 8760
gtgatattat cccattccat gcggggtatc gtatgcttcc ttcagcacta ccctttagct 8820
gttctatatg ctgccactcc tcaattggat tagtctcatc cttcaatgct atcatttcct 8880
ttgatattgg atcatatgca tagtaccgag aaactagtgc gaagtagtga tcaggtattg 8940
ctgttatctg atgagtatac gttgtcctgg ccacggcaga agcacgctta tcgctccaat 9000
ttcccacaac attagtcaac tccgttaggc ccttcattga aagaaatgag gtcatcaaat 9060
gtcttccaat gtgagatttt gggccatttt ttatagcaaa gattgaataa ggcgcatttt 9120
tcttcaaagc tttattgtac gatctgacta agttatcttt taataattgg tattcctgtt 9180
tattgcttga agaattgccg gtcctattta ctcgttttag gactggttca gaattcatcg 9240
atgctcactc aaaggtcggt aatacggtta tccacagaat caggggataa cgcaggaaag 9300
aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg 9360
ttt 9363

<210> 204

<211> 11
<212> PRT
<213> Artificial sequence

<220>
<223> C11 Cter

<400> 204
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro
1 5 10

<210> 205

<211> 29
<212> PRT
<213> Artificial sequence

<220>
<223> C28 Cter

<400> 205
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu
1 5 10 15

Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly
20 25

<210> 206

<211> 40
<212> PRT
<213> Artificial sequence

<220>
<223> C40 Cter

<400> 206
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu
1 5 10 15

Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala
20 25 30

Leu Asp Ala Val Lys Lys Gly Leu
35 40

<210> 207

<211> 64
<212> PRT
<213> Artificial sequence

<220>
<223> C64 Cter

<400> 207
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu
1 5 10 15

Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala
20 25 30

Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile Lys
35 40 45

Arg Thr Asn Arg Arg Ile Pro Glu Arg Thr Ser His Arg Val Ala Asp
50 55 60

<210> 208

<211> 114
<212> PRT
<213> Artificial sequence

<220>
<223> C115 Cter

<400> 208
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu
1 5 10 15

Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala
20 25 30

Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile Lys
35 40 45

Arg Thr Asn Arg Arg Ile Pro Glu Arg Thr Ser His Arg Val Ala Asp
50 55 60

His Ala Gln Val Val Arg Val Leu Gly Phe Phe Gln Cys His Ser His
65 70 75 80

Pro Ala Gln Ala Phe Asp Asp Ala Met Thr Gln Phe Gly Met Ser Arg
85 90 95

His Gly Leu Leu Gln Leu Phe Arg Arg Val Gly Val Thr Glu Leu Glu
100 105 110

Ala Arg

<210> 209

<211> 173
<212> PRT
<213> Artificial sequence

<220>
<223> C172 Cter

<400> 209
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu
1 5 10 15

Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala
20 25 30

Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile Lys
35 40 45

Arg Thr Asn Arg Arg Ile Pro Glu Arg Thr Ser His Arg Val Ala Asp
50 55 60

His Ala Gln Val Val Arg Val Leu Gly Phe Phe Gln Cys His Ser His
65 70 75 80

Pro Ala Gln Ala Phe Asp Asp Ala Met Thr Gln Phe Gly Met Ser Arg
85 90 95

His Gly Leu Leu Gln Leu Phe Arg Arg Val Gly Val Thr Glu Leu Glu
100 105 110

Ala Arg Ser Gly Thr Leu Pro Pro Ala Ser Gln Arg Trp Asp Arg Ile
115 120 125

Leu Gln Ala Ser Gly Met Lys Arg Ala Lys Pro Ser Pro Thr Ser Thr

130 135 140

Gln Thr Pro Asp Gln Ala Ser Leu His Ala Phe Ala Asp Ser Leu Glu
145 150 155 160

Arg Asp Leu Asp Ala Pro Ser Pro Met His Glu Gly Asp
165 170

<210> 210

<211> 231
<212> PRT
<213> Artificial sequence

<220>
<223> WT Cter

<400> 210
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala Ala Leu
1 5 10 15

Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg Pro Ala
20 25 30

Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala Pro Ala Leu Ile Lys
35 40 45

Arg Thr Asn Arg Arg Ile Pro Glu Arg Thr Ser His Arg Val Ala Asp
50 55 60

His Ala Gln Val Val Arg Val Leu Gly Phe Phe Gln Cys His Ser His
65 70 75 80

Pro Ala Gln Ala Phe Asp Asp Ala Met Thr Gln Phe Gly Met Ser Arg
85 90 95

His Gly Leu Leu Gln Leu Phe Arg Arg Val Gly Val Thr Glu Leu Glu
100 105 110

Ala Arg Ser Gly Thr Leu Pro Pro Ala Ser Gln Arg Trp Asp Arg Ile
115 120 125

Leu Gln Ala Ser Gly Met Lys Arg Ala Lys Pro Ser Pro Thr Ser Thr
130 135 140

Gln Thr Pro Asp Gln Ala Ser Leu His Ala Phe Ala Asp Ser Leu Glu
145 150 155 160

Arg Asp Leu Asp Ala Pro Ser Pro Met His Glu Gly Asp Gln Thr Arg
165 170 175

Ala Ser Ser Arg Lys Arg Ser Arg Ser Asp Arg Ala Val Thr Gly Pro
180 185 190

Ser Ala Gln Gln Ser Phe Glu Val Arg Val Pro Glu Gln Arg Asp Ala
195 200 205

Leu His Leu Pro Leu Ser Trp Arg Val Lys Arg Pro Arg Thr Ser Ile
210 215 220

Gly Gly Gly Leu Pro Asp Pro
225 230

<210> 211

<211> 1094
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS7821

<400> 211
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
20 25 30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
35 40 45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
50 55 60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65 70 75 80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
85 90 95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
100 105 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
115 120 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
130 135 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145 150 155 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
165 170 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala

180 185 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
195 200 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
210 215 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225 230 235 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
245 250 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
260 265 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
275 280 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
290 295 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305 310 315 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
325 330 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
340 345 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
355 360 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
370 375 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385 390 395 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
405 410 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
420 425 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
435 440 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln

450 455 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465 470 475 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
485 490 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
500 505 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
515 520 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
530 535 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545 550 555 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
565 570 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
580 585 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
595 600 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
610 615 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625 630 635 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
645 650 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
660 665 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
675 680 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
690 695 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705 710 715 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln

725 730 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
740 745 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
755 760 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
770 775 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785 790 795 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
805 810 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
820 825 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
835 840 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
850 855 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865 870 875 880

Gly Gly Arg Pro Ala Leu Glu Ser Gly Asp Pro Ile Ser Arg Ser Gln
885 890 895

Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His Lys
900 905 910

Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg
915 920 925

Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe Phe
930 935 940

Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser Arg Lys
945 950 955 960

Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly Val
965 970 975

Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly
980 985 990

Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn Gln Thr Arg Asn

995 1000 1005

Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser Val
1010 1015 1020

Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn Tyr
1025 1030 1035 1040

Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn Gly Ala
1045 1050 1055

Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys Ala
1060 1065 1070

Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly Glu
1075 1080 1085

Ile Asn Phe Ala Ala Asp
1090

<210> 212

<211> 1110
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS7803

<400> 212
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
20 25 30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
35 40 45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
50 55 60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65 70 75 80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
85 90 95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
100 105 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
115 120 125

```
Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
    130                 135                 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145                 150                 155                 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
                165                 170                 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
            180                 185                 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
        195                 200                 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
    210                 215                 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225                 230                 235                 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
                245                 250                 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
            260                 265                 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
        275                 280                 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
    290                 295                 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305                 310                 315                 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
                325                 330                 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            340                 345                 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        355                 360                 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
    370                 375                 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385                 390                 395                 400
```

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
405 410 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
420 425 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
435 440 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
450 455 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465 470 475 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
485 490 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
500 505 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
515 520 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
530 535 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545 550 555 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
565 570 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
580 585 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
595 600 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
610 615 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625 630 635 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
645 650 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
660 665 670

```
Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
        675                 680                 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
    690                 695                 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705                 710                 715                 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                725                 730                 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
            740                 745                 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        755                 760                 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
    770                 775                 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785                 790                 795                 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                805                 810                 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            820                 825                 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
        835                 840                 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    850                 855                 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865                 870                 875                 880

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
                885                 890                 895

Asp Pro Ser Gly Ser Gly Ser Gly Gly Asp Pro Ile Ser Arg Ser Gln
            900                 905                 910

Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His Lys
        915                 920                 925

Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg
    930                 935                 940
```

Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe Phe
945 950 955 960

Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser Arg Lys
965 970 975

Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly Val
980 985 990

Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly
995 1000 1005

Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn Gln Thr Arg Asn
1010 1015 1020

Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser Val
1025 1030 1035 1040

Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn Tyr
1045 1050 1055

Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn Gly Ala
1060 1065 1070

Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys Ala
1075 1080 1085

Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly Glu
1090 1095 1100

Ile Asn Phe Ala Ala Asp
1105 1110

<210> 213

<211> 1121
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS7807

<400> 213
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
20 25 30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
35 40 45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala

50 55 60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65 70 75 80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
85 90 95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
100 105 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
115 120 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
130 135 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145 150 155 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
165 170 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
180 185 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
195 200 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
210 215 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225 230 235 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
245 250 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
260 265 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
275 280 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
290 295 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305 310 315 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly

325 330 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
340 345 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
355 360 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
370 375 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385 390 395 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
405 410 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
420 425 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
435 440 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
450 455 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465 470 475 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
485 490 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
500 505 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
515 520 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
530 535 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545 550 555 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
565 570 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
580 585 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly

595 600 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
610 615 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625 630 635 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
645 650 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
660 665 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
675 680 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
690 695 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705 710 715 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
725 730 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
740 745 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
755 760 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
770 775 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785 790 795 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
805 810 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
820 825 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
835 840 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
850 855 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly

865 870 875 880

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
885 890 895

Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala
900 905 910

Cys Leu Gly Gly Asp Pro Ile Ser Arg Ser Gln Leu Val Lys Ser Glu
915 920 925

Leu Glu Glu Lys Lys Ser Glu Leu Arg His Lys Leu Lys Tyr Val Pro
930 935 940

His Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg Asn Ser Thr Gln Asp
945 950 955 960

Arg Ile Leu Glu Met Lys Val Met Glu Phe Phe Met Lys Val Tyr Gly
965 970 975

Tyr Arg Gly Lys His Leu Gly Gly Ser Arg Lys Pro Asp Gly Ala Ile
980 985 990

Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly Val Ile Val Asp Thr Lys
995 1000 1005

Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly Gln Ala Asp Glu Met
1010 1015 1020

Gln Arg Tyr Val Glu Glu Asn Gln Thr Arg Asn Lys His Ile Asn Pro
1025 1030 1035 1040

Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser Val Thr Glu Phe Lys Phe
1045 1050 1055

Leu Phe Val Ser Gly His Phe Lys Gly Asn Tyr Lys Ala Gln Leu Thr
1060 1065 1070

Arg Leu Asn His Ile Thr Asn Cys Asn Gly Ala Val Leu Ser Val Glu
1075 1080 1085

Glu Leu Leu Ile Gly Gly Glu Met Ile Lys Ala Gly Thr Leu Thr Leu
1090 1095 1100

Glu Glu Val Arg Arg Lys Phe Asn Asn Gly Glu Ile Asn Phe Ala Ala
1105 1110 1115 1120

Asp

<210> 214

<211> 1133
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS7809

<400> 214
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
20 25 30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
35 40 45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
50 55 60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65 70 75 80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
85 90 95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
100 105 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
115 120 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
130 135 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145 150 155 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
165 170 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
180 185 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
195 200 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
210 215 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225 230 235 240

```
Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
                245                 250                 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
            260                 265                 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
        275                 280                 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
    290                 295                 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305                 310                 315                 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
                325                 330                 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            340                 345                 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        355                 360                 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
    370                 375                 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385                 390                 395                 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                405                 410                 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            420                 425                 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
        435                 440                 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
    450                 455                 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465                 470                 475                 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                485                 490                 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
            500                 505                 510
```

```
Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        515                 520                 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    530                 535                 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545                 550                 555                 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
                565                 570                 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            580                 585                 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
        595                 600                 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
    610                 615                 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625                 630                 635                 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
                645                 650                 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
            660                 665                 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
        675                 680                 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
    690                 695                 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705                 710                 715                 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                725                 730                 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
            740                 745                 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        755                 760                 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
    770                 775                 780
```

```
Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785                 790                 795                 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                805                 810                 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            820                 825                 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
        835                 840                 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    850                 855                 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865                 870                 875                 880

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
                885                 890                 895

Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala
            900                 905                 910

Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Gly
        915                 920                 925

Asp Pro Ile Ser Arg Ser Gln Leu Val Lys Ser Glu Leu Glu Glu Lys
    930                 935                 940

Lys Ser Glu Leu Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile
945                 950                 955                 960

Glu Leu Ile Glu Ile Ala Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu
                965                 970                 975

Met Lys Val Met Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys
            980                 985                 990

His Leu Gly Gly Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly
        995                 1000                1005

Ser Pro Ile Asp Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly
    1010                1015                1020

Gly Tyr Asn Leu Pro Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val
1025                1030                1035                1040

Glu Glu Asn Gln Thr Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp
                1045                1050                1055
```

Lys Val Tyr Pro Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser
1060 1065 1070

Gly His Phe Lys Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His
1075 1080 1085

Ile Thr Asn Cys Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile
1090 1095 1100

Gly Gly Glu Met Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg
1105 1110 1115 1120

Arg Lys Phe Asn Asn Gly Glu Ile Asn Phe Ala Ala Asp
1125 1130

<210> 215

<211> 1157
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS7811

<400> 215
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
20 25 30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
35 40 45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
50 55 60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65 70 75 80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
85 90 95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
100 105 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
115 120 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
130 135 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly

145 150 155 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
165 170 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
180 185 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
195 200 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
210 215 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225 230 235 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
245 250 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
260 265 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
275 280 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
290 295 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305 310 315 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
325 330 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
340 345 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
355 360 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
370 375 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385 390 395 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
405 410 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val

420 425 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
435 440 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
450 455 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465 470 475 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
485 490 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
500 505 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
515 520 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
530 535 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545 550 555 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
565 570 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
580 585 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
595 600 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
610 615 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625 630 635 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
645 650 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
660 665 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
675 680 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val

690 695 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705 710 715 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
725 730 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
740 745 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
755 760 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
770 775 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785 790 795 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
805 810 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
820 825 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
835 840 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
850 855 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865 870 875 880

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
885 890 895

Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala
900 905 910

Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro
915 920 925

His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg
930 935 940

Thr Ser His Arg Val Ala Asp Gly Asp Pro Ile Ser Arg Ser Gln Leu
945 950 955 960

Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg His Lys Leu

965 970 975

Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg Asn
980 985 990

Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe Phe Met
995 1000 1005

Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser Arg Lys Pro
1010 1015 1020

Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly Val Ile
1025 1030 1035 1040

Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly Gln
1045 1050 1055

Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn Gln Thr Arg Asn Lys
1060 1065 1070

His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser Val Thr
1075 1080 1085

Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn Tyr Lys
1090 1095 1100

Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn Gly Ala Val
1105 1110 1115 1120

Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys Ala Gly
1125 1130 1135

Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly Glu Ile
1140 1145 1150

Asn Phe Ala Ala Asp
1155

<210> 216

<211> 1208
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS7813

<400> 216
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
20 25 30

```
Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
        35                  40                  45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
    50                  55                  60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65                  70                  75                  80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
                85                  90                  95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
            100                 105                 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
        115                 120                 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
    130                 135                 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145                 150                 155                 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
                165                 170                 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
            180                 185                 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
        195                 200                 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
    210                 215                 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225                 230                 235                 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
                245                 250                 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
            260                 265                 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
        275                 280                 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
    290                 295                 300
```

```
Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305                 310                 315                 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
                325                 330                 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            340                 345                 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        355                 360                 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
    370                 375                 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385                 390                 395                 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                405                 410                 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            420                 425                 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
        435                 440                 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
    450                 455                 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465                 470                 475                 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                485                 490                 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
            500                 505                 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        515                 520                 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    530                 535                 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545                 550                 555                 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
                565                 570                 575
```

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
580 585 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
595 600 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
610 615 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625 630 635 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
645 650 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
660 665 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
675 680 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
690 695 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705 710 715 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
725 730 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
740 745 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
755 760 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
770 775 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785 790 795 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
805 810 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
820 825 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
835 840 845

```
Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    850                 855                 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865                 870                 875                 880

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
                885                 890                 895

Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala
            900                 905                 910

Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro
        915                 920                 925

His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg
    930                 935                 940

Thr Ser His Arg Val Ala Asp His Ala Gln Val Val Arg Val Leu Gly
945                 950                 955                 960

Phe Phe Gln Cys His Ser His Pro Ala Gln Ala Phe Asp Asp Ala Met
                965                 970                 975

Thr Gln Phe Gly Met Ser Arg His Gly Leu Leu Gln Leu Phe Arg Arg
            980                 985                 990

Val Gly Val Thr Glu Leu Glu Ala Arg Thr Gly Asp Pro Ile Ser Arg
        995                 1000                1005

Ser Gln Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg
    1010                1015                1020

His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile
1025                1030                1035                1040

Ala Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu
                1045                1050                1055

Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser
            1060                1065                1070

Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr
        1075                1080                1085

Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro
    1090                1095                1100

Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn Gln Thr
1105                1110                1115                1120
```

Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser
1125 1130 1135

Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly
1140 1145 1150

Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn
1155 1160 1165

Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile
1170 1175 1180

Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn
1185 1190 1195 1200

Gly Glu Ile Asn Phe Ala Ala Asp
1205

<210> 217

<211> 1264
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS7817

<400> 217
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
20 25 30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
35 40 45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
50 55 60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65 70 75 80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
85 90 95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
100 105 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
115 120 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln

130 135 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145 150 155 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
165 170 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
180 185 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
195 200 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
210 215 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225 230 235 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
245 250 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
260 265 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
275 280 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
290 295 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305 310 315 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
325 330 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
340 345 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
355 360 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
370 375 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385 390 395 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu

405 410 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
420 425 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
435 440 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
450 455 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465 470 475 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
485 490 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
500 505 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
515 520 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
530 535 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545 550 555 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
565 570 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
580 585 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
595 600 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
610 615 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625 630 635 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
645 650 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
660 665 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu

675 680 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
690 695 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705 710 715 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
725 730 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
740 745 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
755 760 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
770 775 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785 790 795 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
805 810 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
820 825 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
835 840 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
850 855 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865 870 875 880

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
885 890 895

Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala
900 905 910

Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro
915 920 925

His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg
930 935 940

Thr Ser His Arg Val Ala Asp His Ala Gln Val Val Arg Val Leu Gly

945 950 955 960

Phe Phe Gln Cys His Ser His Pro Ala Gln Ala Phe Asp Asp Ala Met
965 970 975

Thr Gln Phe Gly Met Ser Arg His Gly Leu Leu Gln Leu Phe Arg Arg
980 985 990

Val Gly Val Thr Glu Leu Glu Ala Arg Ser Gly Thr Leu Pro Pro Ala
995 1000 1005

Ser Gln Arg Trp Asp Arg Ile Leu Gln Ala Ser Gly Met Lys Arg Ala
1010 1015 1020

Lys Pro Ser Pro Thr Ser Thr Gln Thr Pro Asp Gln Ala Ser Leu His
1025 1030 1035 1040

Ala Phe Ala Asp Ser Leu Glu Arg Asp Leu Asp Ala Pro Ser Pro Met
1045 1050 1055

His Glu Gly Asp Pro Ile Ser Arg Ser Gln Leu Val Lys Ser Glu Leu
1060 1065 1070

Glu Glu Lys Lys Ser Glu Leu Arg His Lys Leu Lys Tyr Val Pro His
1075 1080 1085

Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg Asn Ser Thr Gln Asp Arg
1090 1095 1100

Ile Leu Glu Met Lys Val Met Glu Phe Phe Met Lys Val Tyr Gly Tyr
1105 1110 1115 1120

Arg Gly Lys His Leu Gly Gly Ser Arg Lys Pro Asp Gly Ala Ile Tyr
1125 1130 1135

Thr Val Gly Ser Pro Ile Asp Tyr Gly Val Ile Val Asp Thr Lys Ala
1140 1145 1150

Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly Gln Ala Asp Glu Met Gln
1155 1160 1165

Arg Tyr Val Glu Glu Asn Gln Thr Arg Asn Lys His Ile Asn Pro Asn
1170 1175 1180

Glu Trp Trp Lys Val Tyr Pro Ser Ser Val Thr Glu Phe Lys Phe Leu
1185 1190 1195 1200

Phe Val Ser Gly His Phe Lys Gly Asn Tyr Lys Ala Gln Leu Thr Arg
1205 1210 1215

Leu Asn His Ile Thr Asn Cys Asn Gly Ala Val Leu Ser Val Glu Glu

1220 1225 1230

Leu Leu Ile Gly Gly Glu Met Ile Lys Ala Gly Thr Leu Thr Leu Glu
1235 1240 1245

Glu Val Arg Arg Lys Phe Asn Asn Gly Glu Ile Asn Phe Ala Ala Asp
1250 1255 1260

<210> 218

<211> 5837
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS9012

<400> 218
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 60
tggccttttg ctcacatggt ctttcctgcg ttatcccctg attctgtgga taaccgtatt 120
accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca 180
gtgagcgagg aagcggagag cgcccaatac gcaaaccgcc tctccccgcg cgttggccga 240
ttcattaatg cagctggcac gacaggtttc ccgactggaa agcgggcagt gagcgcaacg 300
caattaatgt gagttagctc actcattagg caccccaggc tttacacttt atgcttccgg 360
ctcgtatgtt gtgtggaatt gtgagcggat aacaatttca cacaggaaac agctatgacc 420
atgattacgc caagcgcgtc aattaaccct cactaaaggg aacaaaagct gttaattaag 480
cgcgccacac gcaaacacaa atacacagcg gccttgccac catggtggat ctacgcacgc 540
tcggctacag ccagcagcaa caggagaaga tcaaaccgaa ggttcgttcg acagtggcgc 600
agcaccacga ggcactggtc ggccacgggt ttacacacgc gcacatcgtt gcgttaagcc 660
aacacccggc agcgttaggg accgtcgctg tcaagtatca ggacatgatc gcagcgttgc 720
cagaggcgac acacgaagcg atcgttggcg tcggcaaaca gtggtccggc gcacgcgctc 780
tggaggcctt gctcacggtg gcgggagagt tgagaggtcc accgttacag ttggacacag 840
gccaacttct caagattgca aaacgtggcg gcgtgaccgc agtggaggca gtgcatgcat 900
ggcgcaatgc actgacgggt gccccgctca acttgacccc ggagcaggtg gtggccatcg 960
ccagccacga tggcggcaag caggcgctgg agacggtcca gcggctgttg ccggtgctgt 1020
gccaggccca cggcttgacc ccccagcagg tggtggccat cgccagcaat ggcggtggca 1080
agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc cacggcttga 1140
ccccggagca ggtggtggcc atcgccagca atattggtgg caagcaggcg ctggagacgg 1200
tgcaggcgct gttgccggtg ctgtgccagg cccacggctt gacccccag caggtggtgg 1260
ccatcgccag caatggcggt ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg 1320
tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc agcaatattg 1380
gtggcaagca ggcgctggag acggtgcagg cgctgttgcc ggtgctgtgc caggcccacg 1440
gcttgacccc ggagcaggtg gtggccatcg ccagcaatat tggtggcaag caggcgctgg 1500
agacggtgca ggcgctgttg ccggtgctgt gccaggccca cggcttgacc ccggagcagg 1560
tggtggccat cgccagcaat attggtggca agcaggcgct ggagacggtg caggcgctgt 1620
tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc atcgccagcc 1680
acgatggcgg caagcaggcg ctggagacgg tccagcggct gttgccggtg ctgtgccagg 1740
cccacggctt gaccccggag caggtggtgg ccatcgccag ccacgatggc ggcaagcagg 1800
cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc ttgaccccc 1860
agcaggtggt ggccatcgcc agcaatggcg gtggcaagca ggcgctggag acggtccagc 1920
ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg gtggccatcg 1980
ccagcaatat tggtggcaag caggcgctgg agacggtgca ggcgctgttg ccggtgctgt 2040
gccaggccca cggcttgacc ccggagcagg tggtggccat cgccagcaat attggtggca 2100
agcaggcgct ggagacggtg caggcgctgt tgccggtgct gtgccaggcc cacggcttga 2160
ccccggagca ggtggtggcc atcgccagcc acgatggcgg caagcaggcg ctggagacgg 2220
tccagcggct gttgccggtg ctgtgccagg cccacggctt gaccccggag caggtggtgg 2280
ccatcgccag ccacgatggc ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg 2340
tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc agccacgatg 2400
gcggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc caggcccacg 2460
gcttgacccc ccagcaggtg gtggccatcg ccagcaatgg cggtggcaag caggcgctgg 2520
agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc ccggagcagg 2580
tggtggccat cgccagccac gatggcggca agcaggcgct ggagacggtc cagcggctgt 2640
tgccggtgct gtgccaggcc cacggcttga cccctcagca ggtggtggcc atcgccagca 2700
atggcggcgg caggccggcg ctggagagca ttgttgccca gttatctcgc cctgatccgg 2760

gatccaagtc tggcatctac cagattaaga acaccctgaa taacaaagtg tatgttggat 2820
ctgccaagga ctttgagaaa agatggaaga ggcacttcaa ggatctggaa aagggttgcc 2880
acagctccat caagctccaa aggtctttca acaaacatgg caatgtcttt gagtgctcca 2940
ttttggaaga gatcccttat gaaaaggacc tcattataga gagggaaaac ttctggatca 3000
aggagctgaa ctcaaaaatc aatggataca acattgcaga tgccacattt ggtgacacct 3060
gttctactca ccccccttaag gaagagatca ttaaaaagag gtctgagact gtgaaggcta 3120
agatgttgaa actggggcct gatggcagaa aggcactcta ctctaagcca ggtagcaaaa 3180
atggaaggtg gaaccctgag acccacaagt tctgcaaatg tggggtgaga atacagacat 3240
cagcctacac ttgcagcaag tgtaggaaca gatcaggaga gaataactcc tttttcaacc 3300
acaagcattc tgacatcacc aagtccaaaa tttcaggcaa gatgaaaggt caaggtccat 3360
cggccgactg ataagtcgac caccaccacc accaccactg ataagagctc ctcgagcctg 3420
cagcagctga agctttggac ttctggcgcg cccaattcgc cctatagtga gtcgtattac 3480
gtcgcgctca ctggccgtcg ttttacaacg tcgtgactgg gaaaaccctg gcgttaccca 3540
acttaatcgc cttgcagcac atccccccttt cgccagctgg cgtaatagcg aagaggcccg 3600
caccgaaacg cccttcccaa cagttgcgca gcctgaatgg cgaatgggag cgccctgtag 3660
cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag 3720
cgccctagcg cccgctcctt tcgctttctt cccttccttt ctcgccacgt tcgccggctt 3780
tccccgtcaa gctctaaatc gggggctccc tttagggttc cgatttagtg ctttacggca 3840
cctcgacccc aaaaaacttg attagggtga tggttggcct gtagtgggcc atagccctga 3900
tagacggttt ttcgcccttt gacgttggag tccacgttct ttaatagtgg actcttgttc 3960
caaactggaa caacactcaa ccctatctcg gtctattctt ttgatttata agggattttg 4020
ccgatttcgg cctattggtt aaaaaatgag ctgatttaac aaaaatttaa cgcgaatttt 4080
aacaaaatat taacgcttac aatttaggtg gcacttttcg gggaaatgtg cgcggaaccc 4140
ctatttgttt atttttctaa atacattcaa atatgtatcc gctcatgaga caataaccct 4200
gataaatgct tcaataatat tgaaaaagga agagtatgag tattcaacat ttccgtgtcg 4260
cccttattcc ctttttgcg gcattttgcc ttcctgtttt tgctcaccca gaaacgctgg 4320
tgaaagtaaa agatgctgaa gatcagttgg gtgcacgagt gggttacatc gaactggatc 4380
tcaacagcgg taagatcctt gagagttttc gccccgaaga acgttttcca atgatgagca 4440
cttttaaagt tctgctatgt ggcgcggtat tatcccgtat tgacgccggg caagagcaac 4500
tcggtcgccg catacactat tctcagaatg acttggttga gtactcacca gtcacagaaa 4560
agcatcttac ggatggcatg acagtaagag aattatgcag tgctgccata accatgagtg 4620
ataacactgc ggccaactta cttctgacaa cgatcggagg accgaaggag ctaaccgctt 4680
ttttgcacaa catgggggat catgtaactc gccttgatcg ttgggaaccg gagctgaatg 4740
aagccatacc aaacgacgag cgtgacacca cgatgcctgt agcaatggca acaacgttgc 4800
gcaaactatt aactggcgaa ctacttactc tagcttcccg gcaacaatta atagactgga 4860
tggaggcgga taaagttgca ggaccacttc tgcgctcggc ccttccggct ggctggttta 4920
ttgctgataa atctggagcc ggtgagcgtg gttctcgcgg tatcattgca gcactggggc 4980
cagatggtaa gccctcccgt atcgtagtta tctacacgac ggggagtcag gcaactatgg 5040
atgaacgaaa tagacagatc gctgagatag gtgcctcact gattaagcat tggtaactgt 5100
cagaccaagt ttactcatat atactttaga ttgatttaaa acttcatttt taatttaaaa 5160
ggatctaggt gaagatcctt tttgataatc tcatgaccaa aatcccttaa cgtgagtttt 5220
cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga gatccttttt 5280
ttctgcgcgt aatctgctgc ttgcaaacaa aaaaaccacc gctaccagcg gtggtttgtt 5340
tgccggatca agagctacca actcttttc cgaaggtaac tggcttcagc agagcgcaga 5400
taccaaatac tgttcttcta gtgtagccgt agttaggcca ccacttcaag aactctgtag 5460
caccgcctac atacctcgct ctgctaatcc tgttaccagt ggctgctgcc agtggcgata 5520
agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg cagcggtcgg 5580
gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac accgaactga 5640
gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca 5700
ggtatccggt aagcggcagg gtcggaacag gagagcgcac gagggagctt ccagggggaa 5760
acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag cgtcgatttt 5820
tgtgatgctc gtcaggg 5837

<210> 219

<211> 6
<212> PRT
<213> Artificial sequence

<220>
<223> SGGSGS stretch

<400> 219
Ser Gly Gly Ser Gly Ser
1 5

<210> 220

<211> 1164
<212> PRT
<213> Xanthomonas spp.

<220>
<223> Original AvrBs3 TALE

<400> 220

```
Met Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu Leu
1               5                  10                  15

Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly Val
            20                  25                  30

Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg Thr
        35                  40                  45

Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala Phe
    50                  55                  60

Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser Leu
65                  70                  75                  80

Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His His
                85                  90                  95

Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu Arg
            100                 105                 110

Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala Ala
        115                 120                 125

Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln Pro
    130                 135                 140

Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr
145                 150                 155                 160

Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val
                165                 170                 175

Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His
            180                 185                 190

Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val
        195                 200                 205

Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala
    210                 215                 220

Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala
225                 230                 235                 240
```

```
Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp
                245                 250                 255

Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val
            260                 265                 270

Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn
        275                 280                 285

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
    290                 295                 300

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
305                 310                 315                 320

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
                325                 330                 335

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
            340                 345                 350

Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
        355                 360                 365

Ser Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
    370                 375                 380

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
385                 390                 395                 400

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
                405                 410                 415

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
            420                 425                 430

Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala
        435                 440                 445

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
    450                 455                 460

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val
465                 470                 475                 480

Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
                485                 490                 495

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu
            500                 505                 510
```

Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
515 520 525

Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala
530 535 540

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
545 550 555 560

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
565 570 575

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
580 585 590

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
595 600 605

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
610 615 620

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn
625 630 635 640

Ser Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val
645 650 655

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
660 665 670

Ser Asn Ser Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
675 680 685

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
690 695 700

Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
705 710 715 720

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
725 730 735

Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val
740 745 750

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
755 760 765

Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu
770 775 780

```
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
785                 790                 795                 800

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala
                805                 810                 815

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
            820                 825                 830

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
        835                 840                 845

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
    850                 855                 860

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
865                 870                 875                 880

Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp
                885                 890                 895

Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala Cys
            900                 905                 910

Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro His
        915                 920                 925

Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg Thr
    930                 935                 940

Ser His Arg Val Ala Asp His Ala Gln Val Val Arg Val Leu Gly Phe
945                 950                 955                 960

Phe Gln Cys His Ser His Pro Ala Gln Ala Phe Asp Asp Ala Met Thr
                965                 970                 975

Gln Phe Gly Met Ser Arg His Gly Leu Leu Gln Leu Phe Arg Arg Val
            980                 985                 990

Gly Val Thr Glu Leu Glu Ala Arg Ser Gly Thr Leu Pro Pro Ala Ser
        995                 1000                1005

Gln Arg Trp Asp Arg Ile Leu Gln Ala Ser Gly Met Lys Arg Ala Lys
    1010                1015                1020

Pro Ser Pro Thr Ser Thr Gln Thr Pro Asp Gln Ala Ser Leu His Ala
1025                1030                1035                1040

Phe Ala Asp Ser Leu Glu Arg Asp Leu Asp Ala Pro Ser Pro Met His
                1045                1050                1055
```

Glu Gly Asp Gln Thr Arg Ala Ser Ser Arg Lys Arg Ser Arg Ser Asp
1060 1065 1070

Arg Ala Val Thr Gly Pro Ser Ala Gln Gln Ser Phe Glu Val Arg Val
1075 1080 1085

Pro Glu Gln Arg Asp Ala Leu His Leu Pro Leu Ser Trp Arg Val Lys
1090 1095 1100

Arg Pro Arg Thr Ser Ile Gly Gly Gly Leu Pro Asp Pro Gly Thr Pro
1105 1110 1115 1120

Thr Ala Ala Asp Leu Ala Ala Ser Ser Thr Val Met Arg Glu Gln Asp
1125 1130 1135

Glu Asp Pro Phe Ala Gly Ala Ala Asp Asp Phe Pro Ala Phe Asn Glu
1140 1145 1150

Glu Glu Leu Ala Trp Leu Met Glu Leu Leu Pro Gln
1155 1160

<210> 221

<211> 417
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS9937

<400> 221
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
100 105 110

```
Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly
    130                 135                 140

Ile Arg Ser Arg Ala Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro
145                 150                 155                 160

Asp Pro Gly Ser Gln Val Pro Pro Leu Thr Glu Leu Ser Pro Ser Ile
                165                 170                 175

Ser Val His Leu Leu Leu Gly Asn Pro Ser Gly Ala Thr Pro Thr Lys
            180                 185                 190

Leu Thr Pro Asp Asn Tyr Leu Met Val Lys Asn Gln Tyr Ala Leu Ser
        195                 200                 205

Tyr Asn Asn Ser Lys Gly Thr Ala Asn Trp Val Ala Trp Gln Leu Asn
    210                 215                 220

Ser Ser Trp Leu Gly Asn Ala Glu Arg Gln Asp Asn Phe Arg Pro Asp
225                 230                 235                 240

Lys Thr Leu Pro Ala Gly Trp Val Arg Val Thr Pro Ser Met Tyr Ser
                245                 250                 255

Gly Ser Gly Tyr Asp Arg Gly His Ile Ala Pro Ser Ala Asp Arg Thr
            260                 265                 270

Lys Thr Thr Glu Asp Asn Ala Ala Thr Phe Leu Met Thr Asn Met Met
        275                 280                 285

Pro Gln Thr Pro Asp Asn Asn Arg Asn Thr Trp Gly Asn Leu Glu Asp
    290                 295                 300

Tyr Cys Arg Glu Leu Val Ser Gln Gly Lys Glu Leu Tyr Ile Val Ala
305                 310                 315                 320

Gly Pro Asn Gly Ser Leu Gly Lys Pro Leu Lys Gly Lys Val Thr Val
                325                 330                 335

Pro Lys Ser Thr Trp Lys Ile Val Val Val Leu Asp Ser Pro Gly Ser
            340                 345                 350

Gly Leu Glu Gly Ile Thr Ala Asn Thr Arg Val Ile Ala Val Asn Ile
        355                 360                 365

Pro Asn Asp Pro Glu Leu Asn Asn Asp Trp Arg Ala Tyr Lys Val Ser
    370                 375                 380
```

Val Asp Glu Leu Glu Ser Leu Thr Gly Tyr Asp Phe Leu Ser Asn Val
385 390 395 400

Ser Pro Asn Ile Gln Thr Ser Ile Glu Ser Lys Val Asp Asn Ala Ala
405 410 415

Asp

<210> 222

<211> 1001
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS9938

<400> 222
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
100 105 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
115 120 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val
130 135 140

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145 150 155 160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
165 170 175

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
180 185 190

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
195 200 205

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
210 215 220

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225 230 235 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
245 250 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
260 265 270

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
275 280 285

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
290 295 300

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile
305 310 315 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu
325 330 335

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
340 345 350

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
355 360 365

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
370 375 380

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
385 390 395 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
405 410 415

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
420 425 430

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
435 440 445

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
450 455 460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465 470 475 480

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
485 490 495

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
500 505 510

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
515 520 525

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
530 535 540

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
545 550 555 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
565 570 575

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
580 585 590

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
595 600 605

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
610 615 620

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
625 630 635 640

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
645 650 655

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
660 665 670

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
675 680 685

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
690 695 700

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
705 710 715 720

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser
725 730 735

```
Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Gly Ser Gln Val Pro Pro
            740                 745                 750

Leu Thr Glu Leu Ser Pro Ser Ile Ser Val His Leu Leu Leu Gly Asn
        755                 760                 765

Pro Ser Gly Ala Thr Pro Thr Lys Leu Thr Pro Asp Asn Tyr Leu Met
    770                 775                 780

Val Lys Asn Gln Tyr Ala Leu Ser Tyr Asn Asn Ser Lys Gly Thr Ala
785                 790                 795                 800

Asn Trp Val Ala Trp Gln Leu Asn Ser Ser Trp Leu Gly Asn Ala Glu
                805                 810                 815

Arg Gln Asp Asn Phe Arg Pro Asp Lys Thr Leu Pro Ala Gly Trp Val
            820                 825                 830

Arg Val Thr Pro Ser Met Tyr Ser Gly Ser Gly Tyr Asp Arg Gly His
        835                 840                 845

Ile Ala Pro Ser Ala Asp Arg Thr Lys Thr Thr Glu Asp Asn Ala Ala
    850                 855                 860

Thr Phe Leu Met Thr Asn Met Met Pro Gln Thr Pro Asp Asn Asn Arg
865                 870                 875                 880

Asn Thr Trp Gly Asn Leu Glu Asp Tyr Cys Arg Glu Leu Val Ser Gln
                885                 890                 895

Gly Lys Glu Leu Tyr Ile Val Ala Gly Pro Asn Gly Ser Leu Gly Lys
            900                 905                 910

Pro Leu Lys Gly Lys Val Thr Val Pro Lys Ser Thr Trp Lys Ile Val
        915                 920                 925

Val Val Leu Asp Ser Pro Gly Ser Gly Leu Glu Gly Ile Thr Ala Asn
    930                 935                 940

Thr Arg Val Ile Ala Val Asn Ile Pro Asn Asp Pro Glu Leu Asn Asn
945                 950                 955                 960

Asp Trp Arg Ala Tyr Lys Val Ser Val Asp Glu Leu Glu Ser Leu Thr
                965                 970                 975

Gly Tyr Asp Phe Leu Ser Asn Val Ser Pro Asn Ile Gln Thr Ser Ile
            980                 985                 990

Glu Ser Lys Val Asp Asn Ala Ala Asp
        995                 1000
```

<210> 223

<211> 11304
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS9924

<400> 223
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 60
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 120
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 180
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 240
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 300
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 360
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 420
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 480
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 540
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 600
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 660
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 720
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 780
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 840
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 900
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 960
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1020
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1080
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1140
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1200
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1260
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1320
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1380
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1440
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1500
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1560
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1620
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1680
atatttgaat gtatttagaa aaataaacaa atagggttc cgcgcacatt tccccgaaaa 1740
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1800
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1860
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1920
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 1980
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2040
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2100
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2160
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2220
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2280
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2340
gacacgcaaa cacaaataca cagcggcctt gccaccatgg tggatctacg cacgctcggc 2400
tacagccagc agcaacagga gaagatcaaa ccgaaggttc gttcgacagt ggcgcagcac 2460
cacgaggcac tggtcggcca cgggtttaca cacgcgcaca tcgttgcgtt aagccaacac 2520
ccggcagcgt tagggaccgt cgctgtcaag tatcaggaca tgatcgcagc gttgccagag 2580
gcgacacacg aagcgatcgt tggcgtcggc aaacagtggt ccggcgcacg cgctctggag 2640
gccttgctca cggtggcggg agagttgaga ggtccaccgt tacagttgga cacaggccaa 2700
cttctcaaga ttgcaaaacg tggcggcgtg accgcagtgg aggcagtgca tgcatggcgc 2760
aatgcactga cgggtgcccc gctcaacttg accccggagc aggtggtggc catcgccagc 2820
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2880
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 2940
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 3000
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 3060
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 3120
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3180
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 3240
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 3300
accccggagc aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg 3360
gtgcaggcgc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 3420

```
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg   3480
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat   3540
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac   3600
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg   3660
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag   3720
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg   3780
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc   3840
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag   3900
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag   3960
gcgctggaga cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg   4020
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag   4080
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc   4140
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg   4200
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc   4260
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg   4320
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg   4380
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg   4440
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg   4500
gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg tggccatcgc cagcaatggc   4560
ggcggcaggc cggcgctgga gagcattgtt gcccagttat ctcgccctga tccgggatcc   4620
caagtcccac cattgacaga gttaagtccg tctatttcgg tacacttact gcttgggaat   4680
ccttctggtg caacacccac aaagctaact ccagacaatt acctgatggt gaagaaccag   4740
tatgccttgt cctacaataa ctccaagggt actgccaatt gggttgcttg gcaattgaac   4800
tcctcttggc ttggtaatgc ggaaagacag gacaatttca gacctgacaa aaccttacct   4860
gcaggatggg taagagttac tccatcaatg tattccggat ctggttatga tagaggtcat   4920
atagctccta gtgctgatag gacgaaaact actgaggata atgccgctac ctttctgatg   4980
acgaacatga tgcctcaaac tccagataat aaccgtaata cctggggaaa cttggaagat   5040
tactgtaggg aattagtctc acaaggtaaa gaattgtaca ttgttgcagg acctaatggc   5100
tcattaggca aacccttgaa agggaaagtc actgtgccaa agagtacatg gaagattgtg   5160
gttgttcttg attcaccagg tagtggtttg gaaggtataa cggcaaacac aagagtgatt   5220
gctgtcaata tcccaaacga tccggagttg aacaacgatt ggagagcgta taaggtatca   5280
gttgacgaac tagaatcttt aacaggctat gactttctaa gcaatgtatc tcccaatatc   5340
caaaccagca tagaatcgaa agttgataat gcggccgact gataactcga gcgatcctct   5400
agacgagctc ctcgagcctg cagcagctga agctttggac ttcttcgcca gaggtttggt   5460
caagtctcca atcaaggttg tcggcttgtc taccttgcca gaaatttacg aaaagatgga   5520
aaagggtcaa atcgttggta gatacgttgt tgacacttct aaataagcga atttcttatg   5580
atttatgatt tttattatta aataagttat aaaaaaaata agtgtataca aattttaaag   5640
tgactcttag gttttaaaac gaaaattctt attcttgagt aactctttcc tgtaggtcag   5700
gttgctttct caggtatagc atgaggtcgc tcttattgac cacacctcta ccggcatgca   5760
agcttggcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc gctcacaatt   5820
ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc   5880
taactcacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc   5940
cagcagatct attacattat gggtggtatg ttggaataaa aatcaactat catctactaa   6000
ctagtattta cgttactagt atattatcat atacggtgtt agaagatgac gcaaatgatg   6060
agaaatagtc atctaaatta gtggaagctg aaacgcaagg attgataatg taataggatc   6120
aatgaatatt aacatataaa atgatgataa taatatttat agaattgtgt agaattgcag   6180
attccctttt atggattcct aaatcctcga ggagaacttc tagtatatct acatacctaa   6240
tattattgcc ttattaaaaa tggaatccca acaattacat caaaatccac attctcttca   6300
aaatcaattg tcctgtactt ccttgttcat gtgtgttcaa aaacgttata tttataggat   6360
aattatactc tatttctcaa caagtaattg gttgtttggc cgagcggtct aaggcgcctg   6420
attcaagaaa tatcttgacc gcagttaact gtgggaatac tcaggtatcg taagatgcaa   6480
gagttcgaat ctcttagcaa ccattatttt tttcctcaac ataacgagaa cacacagggg   6540
cgctatcgca cagaatcaaa ttcgatgact ggaaattttt tgttaatttc agaggtcgcc   6600
tgacgcatat acctttttca actgaaaaat tgggagaaaa aggaaaggtg agagccgcgg   6660
aaccggcttt tcatatagaa tagagaagcg ttcatgacta aatgcttgca tcacaatact   6720
tgaagttgac aatattattt aaggacctat tgtttttttcc aataggtggt tagcaatcgt   6780
cttactttct aacttttctt acctttttaca tttcagcaat atatatatat atatttcaag   6840
gatataccat tctaatgtct gcccctaaga agatcgtcgt tttgccaggt gaccacgttg   6900
gtcaagaaat cacagccgaa gccattaagg ttcttaaagc tatttctgat gttcgttcca   6960
atgtcaagtt cgatttcgaa aatcatttaa ttggtggtgc tgctatcgat gctacaggtg   7020
tcccacttcc agatgaggcg ctggaagcct ccaagaaggt tgatgccgtt ttgttaggtg   7080
ctgtgggtgg tcctaaatgg ggtaccggta gtgttagacc tgaacaaggt ttactaaaaa   7140
tccgtaaaga acttcaattg tacgccaact taagaccatg taactttgca tccgactctc   7200
ttttagactt atctccaatc aagccacaat ttgctaaagg tactgacttc gttgttgtca   7260
gagaattagt gggaggtatt tactttggta agagaaagga agacgatggt gatggtgtcg   7320
cttgggatag tgaacaatac accgttccag aagtgcaaag aatcacaaga atggccgctt   7380
tcatggccct acaacatgag ccaccattgc ctatttggtc cttggataaa gctaatgttt   7440
tggcctcttc aagattatgg agaaaaactg tggaggaaac catcaagaac gaattcccta   7500
```

cattgaaggt tcaacatcaa ttgattgatt ctgccgccat gatcctagtt aagaacccaa 7560
cccacctaaa tggtattata atcaccagca acatgtttgg tgatatcatc tccgatgaag 7620
cctccgttat cccaggttcc ttgggtttgt tgccatctgc gtccttggcc tctttgccag 7680
acaagaacac cgcatttggt ttgtacgaac catgccacgg ttctgctcca gatttgccaa 7740
agaataaggt caaccctatc gccactatct tgtctgctgc aatgatgttg aaattgtcat 7800
tgaacttgcc tgaagaaggt aaggccattg aagatgcagt taaaaaggtt ttggatgcag 7860
gtatcagaac tggtgattta ggtggttcca acagtaccac ggaagtcggt gatgctgtcg 7920
ccgaagaagt taagaaaatc cttgcttaaa aagattctct ttttttatga tatttgtaca 7980
taaactttat aaatgaaatt cataatagaa acgacacgaa attacaaaat ggaatatgtt 8040
catagggtag acgaaactat atacgcaatc tacatacatt tatcaagaag gagaaaaagg 8100
aggatgtaaa ggaatacagg taagcaaatt gatactaatg gctcaacgtg ataaggaaaa 8160
agaattgcac tttaacatta atattgacaa ggaggagggc accacacaaa aagttaggtg 8220
taacagaaaa tcatgaaact atgattccta atttatatat tggaggattt tctctaaaaa 8280
aaaaaaaata caacaaataa aaaacactca atgacctgac catttgatgg agtttaagtc 8340
aataccttct tgaaccattt cccataatgg tgaaagttcc ctcaagaatt ttactctgtc 8400
agaaacggcc ttaacgacgt agtcgacctc ctcttcagta ctaaatctac caataccaaa 8460
tctgatggaa gaatgggcta atgcatcatc cttacccagc gcatgtaaaa cataagaagg 8520
ttctagggaa gcagatgtac aggctgaacc cgaggataat gcgatatccc ttagtgccat 8580
caataaagat tctccttcca cgtaggcgaa agaaacgtta acacaccctg gataacgatg 8640
atctggagat ccgttcaacg tggtatgttc agcggataat agacctttga ctaatttatc 8700
ggatagtctt ttgatgtgag cttggtcgtt gtcaaattct ttcttcatca atctcgcagc 8760
ttcaccaaat cccgctacca atgggggggc caaagtacca gatctgctgc attaatgaat 8820
cggccaacgc gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac 8880
tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagcatcga tgaattccac 8940
ggactataga ctatactagt atactccgtc tactgtacga tacacttccg ctcaggtcct 9000
tgtcctttaa cgaggcctta ccactctttt gttactctat tgatccagct cagcaaaggc 9060
agtgtgatct aagattctat cttcgcgatg tagtaaaact agctagaccg agaaagagac 9120
tagaaatgca aaaggcactt ctacaatggc tgccatcatt attatccgat gtgacgctgc 9180
agcttctcaa tgatattcga atacgctttg aggagataca gcctaatatc cgacaaactg 9240
ttttacagat ttacgatcgt acttgttacc catcattgaa ttttgaacat ccgaacctgg 9300
gagttttccc tgaaacagat agtatatttg aacctgtata ataatatata gtctagcgct 9360
ttacggaaga caatgtatgt atttcggttc ctggagaaac tattgcatct attgcatagg 9420
taatcttgca cgtcgcatcc ccggttcatt ttctgcgttt ccatcttgca cttcaatagc 9480
atatctttgt taacgaagca tctgtgcttc attttgtaga acaaaaatgc aacgcgagag 9540
cgctaatttt tcaaacaaag aatctgagct gcatttttac agaacagaaa tgcaacgcga 9600
aagcgctatt ttaccaacga agaatctgtg cttcattttt gtaaaacaaa aatgcaacgc 9660
gagagcgcta atttttcaaa caaagaatct gagctgcatt tttacagaac agaaatgcaa 9720
cgcgagagcg ctattttacc aacaaagaat ctatacttct tttttgttct acaaaaatgc 9780
atcccgagag cgctattttt ctaacaaagc atcttagatt actttttttc tcctttgtgc 9840
gctctataat gcagtctctt gataactttt tgcactgtag gtccgttaag gttagaagaa 9900
ggctactttg gtgtctattt tctcttccat aaaaaaagcc tgactccact tcccgcgttt 9960
actgattact agcgaagctg cgggtgcatt ttttcaagat aaaggcatcc ccgattatat 10020
tctataccga tgtggattgc gcatactttg tgaacagaaa gtgatagcgt tgatgattct 10080
tcattggtca gaaaattatg aacggtttct tctattttgt ctctatatac tacgtatagg 10140
aaatgtttac attttcgtat tgttttcgat tcactctatg aatagttctt actacaattt 10200
ttttgtctaa agagtaatac tagagataaa cataaaaaat gtagaggtcg agtttagatg 10260
caagttcaag gagcgaaagg tggatgggta ggttatatag ggatatagca cagagatata 10320
tagcaaagag atacttttga gcaatgtttg tggaagcggt attcgcaata ttttagtagc 10380
tcgttacagt ccggtgcgtt tttggttttt tgaaagtgcg tcttcagagc gcttttggtt 10440
ttcaaaagcg ctctgaagtt cctatacttt ctagagaata ggaacttcgg aataggaact 10500
tcaaagcgtt tccgaaaacg agcgcttccg aaaatgcaac gcgagctgcg cacatacagc 10560
tcactgttca cgtcgcacct atatctgcgt gttgcctgta tatatatata catgagaaga 10620
acggcatagt gcgtgtttat gcttaaatgc gtacttatat gcgtctattt atgtaggatg 10680
aaaggtagtc tagtacctcc tgtgatatta tcccattcca tgcggggtat cgtatgcttc 10740
cttcagcact acccttttagc tgttctatat gctgccactc ctcaattgga ttagtctcat 10800
ccttcaatgc tatcatttcc tttgatattg gatcatatgc atagtaccga gaaactagtg 10860
cgaagtagtg atcaggtatt gctgttatct gatgagtata cgttgtcctg gccacggcag 10920
aagcacgctt atcgctccaa tttcccacaa cattagtcaa ctccgttagg cccttcattg 10980
aaagaaatga ggtcatcaaa tgtcttccaa tgtgagattt tgggccattt tttatagcaa 11040
agattgaata aggcgcattt ttcttcaaag ctttattgta cgatctgact aagttatctt 11100
ttaataattg gtattcctgt ttattgcttg aagaattgcc ggtcctattt actcgtttta 11160
ggactggttc agaattcatc gatgctcact caaaggtcgg taatacggtt atccacagaa 11220
tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt 11280
aaaaaggccg cgttgctggc gttt 11304

<210> 224

<211> 65
<212> DNA
<213> Artificial sequence

<220>
<223> compact target

<400> 224
tatataaacc taaccctctt agcatgaagg taccttgtcg atggtgtaca gtaggggggag 60
atgca 65

<210> 225

<211> 1166
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS9596

<400> 225
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
20 25 30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
35 40 45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
50 55 60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65 70 75 80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
85 90 95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
100 105 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
115 120 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
130 135 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145 150 155 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
165 170 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
180 185 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
195 200 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
210 215 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225 230 235 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
245 250 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
260 265 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
275 280 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
290 295 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305 310 315 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
325 330 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
340 345 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
355 360 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
370 375 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385 390 395 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
405 410 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
420 425 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
435 440 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
450 455 460

```
Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465                 470                 475                 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                485                 490                 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
            500                 505                 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        515                 520                 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    530                 535                 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545                 550                 555                 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
                565                 570                 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            580                 585                 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
        595                 600                 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
    610                 615                 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625                 630                 635                 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
                645                 650                 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
            660                 665                 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
        675                 680                 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
    690                 695                 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705                 710                 715                 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                725                 730                 735
```

```
Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
            740                 745                 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        755                 760                 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
    770                 775                 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785                 790                 795                 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                805                 810                 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            820                 825                 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
        835                 840                 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    850                 855                 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865                 870                 875                 880

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
                885                 890                 895

Asp Pro Ser Gly Ser Gly Ser Gly Gly Asp Pro Ser Gly Gly Ser Gly
            900                 905                 910

Ser Gln Val Pro Pro Leu Thr Glu Leu Ser Pro Ser Ile Ser Val His
        915                 920                 925

Leu Leu Leu Gly Asn Pro Ser Gly Ala Thr Pro Thr Lys Leu Thr Pro
    930                 935                 940

Asp Asn Tyr Leu Met Val Lys Asn Gln Tyr Ala Leu Ser Tyr Asn Asn
945                 950                 955                 960

Ser Lys Gly Thr Ala Asn Trp Val Ala Trp Gln Leu Asn Ser Ser Trp
                965                 970                 975

Leu Gly Asn Ala Glu Arg Gln Asp Asn Phe Arg Pro Asp Lys Thr Leu
            980                 985                 990

Pro Ala Gly Trp Val Arg Val Thr Pro Ser Met Tyr Ser Gly Ser Gly
        995                 1000                1005
```

Tyr Asp Arg Gly His Ile Ala Pro Ser Ala Asp Arg Thr Lys Thr Thr
1010 1015 1020

Glu Asp Asn Ala Ala Thr Phe Leu Met Thr Asn Met Met Pro Gln Thr
1025 1030 1035 1040

Pro Asp Asn Asn Arg Asn Thr Trp Gly Asn Leu Glu Asp Tyr Cys Arg
1045 1050 1055

Glu Leu Val Ser Gln Gly Lys Glu Leu Tyr Ile Val Ala Gly Pro Asn
1060 1065 1070

Gly Ser Leu Gly Lys Pro Leu Lys Gly Lys Val Thr Val Pro Lys Ser
1075 1080 1085

Thr Trp Lys Ile Val Val Val Leu Asp Ser Pro Gly Ser Gly Leu Glu
1090 1095 1100

Gly Ile Thr Ala Asn Thr Arg Val Ile Ala Val Asn Ile Pro Asn Asp
1105 1110 1115 1120

Pro Glu Leu Asn Asn Asp Trp Arg Ala Tyr Lys Val Ser Val Asp Glu
1125 1130 1135

Leu Glu Ser Leu Thr Gly Tyr Asp Phe Leu Ser Asn Val Ser Pro Asn
1140 1145 1150

Ile Gln Thr Ser Ile Glu Ser Lys Val Asp Asn Ala Ala Asp
1155 1160 1165

<210> 226

<211> 1177
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS9597

<400> 226
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
20 25 30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
35 40 45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
50 55 60

```
Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65                  70                  75                  80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
                85                  90                  95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
            100                 105                 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
        115                 120                 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
    130                 135                 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145                 150                 155                 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
                165                 170                 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
            180                 185                 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
        195                 200                 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
    210                 215                 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225                 230                 235                 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
                245                 250                 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
            260                 265                 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
        275                 280                 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
    290                 295                 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305                 310                 315                 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
                325                 330                 335
```

```
Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            340                 345                 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        355                 360                 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
    370                 375                 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385                 390                 395                 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                405                 410                 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            420                 425                 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
        435                 440                 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
    450                 455                 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465                 470                 475                 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                485                 490                 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
            500                 505                 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        515                 520                 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    530                 535                 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545                 550                 555                 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
                565                 570                 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            580                 585                 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
        595                 600                 605
```

```
Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
    610                 615                 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625                 630                 635                 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
                645                 650                 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
            660                 665                 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
        675                 680                 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
    690                 695                 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705                 710                 715                 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                725                 730                 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
            740                 745                 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        755                 760                 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
    770                 775                 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785                 790                 795                 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                805                 810                 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            820                 825                 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
        835                 840                 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    850                 855                 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865                 870                 875                 880
```

```
Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
                885                 890                 895

Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala
            900                 905                 910

Cys Leu Gly Gly Asp Pro Ser Gly Gly Ser Gly Ser Gln Val Pro Pro
        915                 920                 925

Leu Thr Glu Leu Ser Pro Ser Ile Ser Val His Leu Leu Leu Gly Asn
    930                 935                 940

Pro Ser Gly Ala Thr Pro Thr Lys Leu Thr Pro Asp Asn Tyr Leu Met
945                 950                 955                 960

Val Lys Asn Gln Tyr Ala Leu Ser Tyr Asn Asn Ser Lys Gly Thr Ala
                965                 970                 975

Asn Trp Val Ala Trp Gln Leu Asn Ser Ser Trp Leu Gly Asn Ala Glu
            980                 985                 990

Arg Gln Asp Asn Phe Arg Pro Asp Lys Thr Leu Pro Ala Gly Trp Val
        995                 1000                1005

Arg Val Thr Pro Ser Met Tyr Ser Gly Ser Gly Tyr Asp Arg Gly His
    1010                1015                1020

Ile Ala Pro Ser Ala Asp Arg Thr Lys Thr Thr Glu Asp Asn Ala Ala
1025                1030                1035                1040

Thr Phe Leu Met Thr Asn Met Met Pro Gln Thr Pro Asp Asn Asn Arg
                1045                1050                1055

Asn Thr Trp Gly Asn Leu Glu Asp Tyr Cys Arg Glu Leu Val Ser Gln
            1060                1065                1070

Gly Lys Glu Leu Tyr Ile Val Ala Gly Pro Asn Gly Ser Leu Gly Lys
        1075                1080                1085

Pro Leu Lys Gly Lys Val Thr Val Pro Lys Ser Thr Trp Lys Ile Val
    1090                1095                1100

Val Val Leu Asp Ser Pro Gly Ser Gly Leu Glu Gly Ile Thr Ala Asn
1105                1110                1115                1120

Thr Arg Val Ile Ala Val Asn Ile Pro Asn Asp Pro Glu Leu Asn Asn
                1125                1130                1135

Asp Trp Arg Ala Tyr Lys Val Ser Val Asp Glu Leu Glu Ser Leu Thr
            1140                1145                1150
```

Gly Tyr Asp Phe Leu Ser Asn Val Ser Pro Asn Ile Gln Thr Ser Ile
1155 1160 1165

Glu Ser Lys Val Asp Asn Ala Ala Asp
1170 1175

<210> 227

<211> 1213
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS9599

<400> 227
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
20 25 30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
35 40 45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
50 55 60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65 70 75 80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
85 90 95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
100 105 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
115 120 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
130 135 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145 150 155 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
165 170 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
180 185 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
195 200 205

```
Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
    210                 215                 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225                 230                 235                 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
                245                 250                 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
            260                 265                 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
        275                 280                 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
    290                 295                 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305                 310                 315                 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
                325                 330                 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            340                 345                 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        355                 360                 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
    370                 375                 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385                 390                 395                 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                405                 410                 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            420                 425                 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
        435                 440                 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
    450                 455                 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465                 470                 475                 480
```

```
Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                485                 490                 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
            500                 505                 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        515                 520                 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    530                 535                 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545                 550                 555                 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
                565                 570                 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            580                 585                 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
        595                 600                 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
    610                 615                 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625                 630                 635                 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
                645                 650                 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
            660                 665                 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
        675                 680                 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
    690                 695                 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705                 710                 715                 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                725                 730                 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
            740                 745                 750
```

```
Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        755                 760                 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
    770                 775                 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785                 790                 795                 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                805                 810                 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            820                 825                 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
        835                 840                 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    850                 855                 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865                 870                 875                 880

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
                885                 890                 895

Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala
            900                 905                 910

Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro
        915                 920                 925

His Ala Pro Ala Leu Ile Lys Arg Thr Asn Arg Arg Ile Pro Glu Arg
    930                 935                 940

Thr Ser His Arg Val Ala Asp Gly Asp Pro Ser Gly Gly Ser Gly Ser
945                 950                 955                 960

Gln Val Pro Pro Leu Thr Glu Leu Ser Pro Ser Ile Ser Val His Leu
                965                 970                 975

Leu Leu Gly Asn Pro Ser Gly Ala Thr Pro Thr Lys Leu Thr Pro Asp
            980                 985                 990

Asn Tyr Leu Met Val Lys Asn Gln Tyr Ala Leu Ser Tyr Asn Asn Ser
        995                 1000                1005

Lys Gly Thr Ala Asn Trp Val Ala Trp Gln Leu Asn Ser Ser Trp Leu
    1010                1015                1020
```

Gly Asn Ala Glu Arg Gln Asp Asn Phe Arg Pro Asp Lys Thr Leu Pro
1025 1030 1035 1040

Ala Gly Trp Val Arg Val Thr Pro Ser Met Tyr Ser Gly Ser Gly Tyr
1045 1050 1055

Asp Arg Gly His Ile Ala Pro Ser Ala Asp Arg Thr Lys Thr Thr Glu
1060 1065 1070

Asp Asn Ala Ala Thr Phe Leu Met Thr Asn Met Met Pro Gln Thr Pro
1075 1080 1085

Asp Asn Asn Arg Asn Thr Trp Gly Asn Leu Glu Asp Tyr Cys Arg Glu
1090 1095 1100

Leu Val Ser Gln Gly Lys Glu Leu Tyr Ile Val Ala Gly Pro Asn Gly
1105 1110 1115 1120

Ser Leu Gly Lys Pro Leu Lys Gly Lys Val Thr Val Pro Lys Ser Thr
1125 1130 1135

Trp Lys Ile Val Val Val Leu Asp Ser Pro Gly Ser Gly Leu Glu Gly
1140 1145 1150

Ile Thr Ala Asn Thr Arg Val Ile Ala Val Asn Ile Pro Asn Asp Pro
1155 1160 1165

Glu Leu Asn Asn Asp Trp Arg Ala Tyr Lys Val Ser Val Asp Glu Leu
1170 1175 1180

Glu Ser Leu Thr Gly Tyr Asp Phe Leu Ser Asn Val Ser Pro Asn Ile
1185 1190 1195 1200

Gln Thr Ser Ile Glu Ser Lys Val Asp Asn Ala Ala Asp
1205 1210

<210> 228

<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> negative control target

<400> 228
ttgttctcag gtacctcagc caga 24

<210> 229

<211> 135
<212> PRT
<213> Artificial sequence

<220>
<223> NuiA protein

<400> 229
Met Thr Lys Thr Asn Ser Glu Ile Leu Glu Gln Leu Lys Gln Ala Ser
1   5   10   15

Asp Gly Leu Leu Phe Met Ser Glu Ser Glu Tyr Pro Phe Glu Val Phe
            20   25   30

Leu Trp Glu Gly Ser Ala Pro Pro Val Thr His Glu Ile Val Leu Gln
        35   40   45

Gln Thr Gly His Gly Gln Asp Ala Pro Phe Lys Val Val Asp Ile Asp
    50   55   60

Ser Phe Phe Ser Arg Ala Thr Thr Pro Gln Asp Trp Tyr Glu Asp Glu
65   70   75   80

Glu Asn Ala Val Val Ala Lys Phe Gln Lys Leu Leu Glu Val Ile Lys
                85   90   95

Ser Asn Leu Lys Asn Pro Gln Val Tyr Arg Leu Gly Glu Val Glu Leu
            100   105   110

Asp Val Tyr Val Ile Gly Glu Thr Pro Ala Gly Asn Leu Ala Gly Ile
        115   120   125

Ser Thr Lys Val Val Glu Thr
    130   135

<210> 230

<211> 87
<212> PRT
<213> Artificial sequence

<220>
<223> Im7 protein

<400> 230
Met Glu Leu Lys Asn Ser Ile Ser Asp Tyr Thr Glu Ala Glu Phe Val
1   5   10   15

Gln Leu Leu Lys Glu Ile Glu Lys Glu Asn Val Ala Ala Thr Asp Asp
            20   25   30

Val Leu Asp Val Leu Leu Glu His Phe Val Lys Ile Thr Glu His Pro
        35   40   45

Asp Gly Thr Asp Leu Ile Tyr Tyr Pro Ser Asp Asn Arg Asp Asp Ser
    50   55   60

Pro Glu Gly Ile Val Lys Glu Ile Lys Glu Trp Arg Ala Ala Asn Gly

65 70 75 80

Lys Pro Gly Phe Lys Gln Gly
85

<210> 231

<211> 303
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS9939

<400> 231
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
100 105 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
115 120 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly
130 135 140

Ile Arg Ser Arg Ala Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro
145 150 155 160

Asp Pro Gly Ser Ser Gly Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr
165 170 175

Gly Lys Gly Lys Pro Val Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys
180 185 190

Asp Leu Gly Ser Pro Val Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp
195 200 205

```
Lys Glu Phe Lys Ser Phe Asp Asp Phe Arg Lys Lys Phe Trp Glu Glu
    210                 215                 220

Val Ser Lys Asp Pro Glu Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn
225                 230                 235                 240

Asp Arg Met Lys Val Gly Lys Ala Pro Lys Thr Arg Thr Gln Asp Val
                245                 250                 255

Ser Gly Lys Arg Thr Ser Phe Glu Leu His His Glu Lys Pro Ile Ser
            260                 265                 270

Gln Asn Gly Gly Val Tyr Asp Met Asp Asn Ile Ser Val Val Thr Pro
        275                 280                 285

Lys Arg His Ile Asp Ile His Arg Gly Lys Gly Ser Ser Ala Asp
    290                 295                 300
```

<210> 232

<211> 887
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS9940

<400> 232

```
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1               5                   10                  15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
            20                  25                  30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
        35                  40                  45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
    50                  55                  60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65                  70                  75                  80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
                85                  90                  95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
            100                 105                 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val
```

130 135 140

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145 150 155 160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
165 170 175

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
180 185 190

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
195 200 205

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
210 215 220

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225 230 235 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
245 250 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
260 265 270

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
275 280 285

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
290 295 300

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile
305 310 315 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu
325 330 335

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
340 345 350

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
355 360 365

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
370 375 380

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
385 390 395 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val

405 410 415

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
420 425 430

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
435 440 445

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
450 455 460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465 470 475 480

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
485 490 495

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
500 505 510

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
515 520 525

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
530 535 540

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
545 550 555 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
565 570 575

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
580 585 590

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
595 600 605

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
610 615 620

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
625 630 635 640

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
645 650 655

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
660 665 670

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val

675 680 685

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
690 695 700

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
705 710 715 720

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser
725 730 735

Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Gly Ser Ser Gly Ser Lys
740 745 750

Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys Gly Lys Pro Val Asn Asn
755 760 765

Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu Gly Ser Pro Val Pro Asp
770 775 780

Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu Phe Lys Ser Phe Asp Asp
785 790 795 800

Phe Arg Lys Lys Phe Trp Glu Glu Val Ser Lys Asp Pro Glu Leu Ser
805 810 815

Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg Met Lys Val Gly Lys Ala
820 825 830

Pro Lys Thr Arg Thr Gln Asp Val Ser Gly Lys Arg Thr Ser Phe Glu
835 840 845

Leu His His Glu Lys Pro Ile Ser Gln Asn Gly Gly Val Tyr Asp Met
850 855 860

Asp Asn Ile Ser Val Val Thr Pro Lys Arg His Ile Asp Ile His Arg
865 870 875 880

Gly Lys Gly Ser Ser Ala Asp
885

<210> 233

<211> 10962
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS8589

<400> 233
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 60
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 120
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 180

gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 240
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 300
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 360
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 420
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 480
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 540
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 600
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 660
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 720
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 780
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 840
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 900
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 960
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1020
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1080
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1140
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1200
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1260
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1320
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1380
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1440
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1500
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1560
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1620
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1680
atatttgaat gtatttagaa aaataaacaa atagggggttc cgcgcacatt tccccgaaaa 1740
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1800
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1860
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1920
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 1980
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2040
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2100
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2160
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2220
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2280
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2340
gacacgcaaa cacaaataca cagcggcctt gccaccatgg tggatctacg cacgctcggc 2400
tacagccagc agcaacagga gaagatcaaa ccgaaggttc gttcgacagt ggcgcagcac 2460
cacgaggcac tggtcggcca cgggtttaca cacgcgcaca tcgttgcgtt aagccaacac 2520
ccggcagcgt tagggaccgt cgctgtcaag tatcaggaca tgatcgcagc gttgccagag 2580
gcgacacacg aagcgatcgt tggcgtcggc aaacagtggt ccggcgcacg cgctctggag 2640
gccttgctca cggtggcggg agagttgaga ggtccaccgt tacagttgga cacaggccaa 2700
cttctcaaga ttgcaaaacg tggcggcgtg accgcagtgg aggcagtgca tgcatggcgc 2760
aatgcactga cgggtgcccc gctcaacttg accccggagc aggtggtggc catcgccagc 2820
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2880
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 2940
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 3000
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 3060
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 3120
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3180
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 3240
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 3300
accccggagc aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg 3360
gtgcaggcgc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 3420
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 3480
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat 3540
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 3600
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg 3660
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 3720
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 3780
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 3840
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 3900
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag 3960
gcgctggaga cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 4020
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 4080
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 4140
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 4200
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 4260

aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 4320
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 4380
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 4440
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 4500
gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg tggccatcgc cagcaatggc 4560
ggcggcaggc cggcgctgga gagcattgtt gcccagttat ctcgccctga tccgggatcc 4620
tccggatcta aacgcaacaa accaggtaag gccactggca aaggtaaacc agtgaacaat 4680
aagtggctga acaacgctgg caaagacttg ggttctccag tgccggatcg cattgccaac 4740
aagctgagag acaaagaatt caagagtttt gatgactttc gcaagaaatt ctgggaagaa 4800
gtgtctaaag atccagaact gtctaaacag tttagtcgca acaacaatga cagaatgaaa 4860
gttggtaaag ctccgaagac ccgcacccaa gacgtgtctg gcaaacgcac cagttttgaa 4920
ttgcatcatg aaaaaccaat ctctcagaat ggtggcgttt acgacatgga caacatctct 4980
gtggttactc cgaaaagaca tatcgatatt catcgcggta aaggatcctc ggccgactga 5040
taactcgagc gatcctctag acgagctcct cgagcctgca gcagctgaag ctttggactt 5100
cttcgccaga ggtttggtca agtctccaat caaggttgtc ggcttgtcta ccttgccaga 5160
aatttacgaa aagatggaaa agggtcaaat cgttggtaga tacgttgttg acacttctaa 5220
ataagcgaat ttcttatgat ttatgatttt tattattaaa taagttataa aaaaaataag 5280
tgtatacaaa ttttaaagtg actcttaggt tttaaaacga aaattcttat tcttgagtaa 5340
ctctttcctg taggtcaggt tgctttctca ggtatagcat gaggtcgctc ttattgacca 5400
cacctctacc ggcatgcaag cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat 5460
tgttatccgc tcacaattcc acacaacata cgagccggaa gcataaagtg taaagcctgg 5520
ggtgcctaat gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag 5580
tcgggaaacc tgtcgtgcca gcagatctat tacattatgg gtggtatgtt ggaataaaaa 5640
tcaactatca tctactaact agtatttacg ttactagtat attatcatat acggtgttag 5700
aagatgacgc aaatgatgag aaatagtcat ctaaattagt ggaagctgaa acgcaaggat 5760
tgataatgta ataggatcaa tgaatattaa catataaaat gatgataata atatttatag 5820
aattgtgtag aattgcagat tccctttttat ggattcctaa atcctcgagg agaacttcta 5880
gtatatctac atacctaata ttattgcctt attaaaaatg gaatcccaac aattacatca 5940
aaatccacat tctcttcaaa atcaattgtc ctgtacttcc ttgttcatgt gtgttcaaaa 6000
acgttatatt tataggataa ttatactcta tttctcaaca agtaattggt tgtttggccg 6060
agcggtctaa ggcgcctgat tcaagaaata tcttgaccgc agttaactgt gggaatactc 6120
aggtatcgta agatgcaaga gttcgaatct cttagcaacc attatttttt tcctcaacat 6180
aacgagaaca cacaggggcg ctatcgcaca gaatcaaatt cgatgactgg aaatttttg 6240
ttaatttcag aggtcgcctg acgcatatac ctttttcaac tgaaaaattg ggagaaaaag 6300
gaaaggtgag agccgcggaa ccggcttttc atatagaata gagaagcgtt catgactaaa 6360
tgcttgcatc acaatacttg aagttgacaa tattatttaa ggacctattg ttttttccaa 6420
taggtggtta gcaatcgtct tactttctaa cttttcttac cttttacatt tcagcaatat 6480
atatatatat atttcaagga tataccattc taatgtctgc ccctaagaag atcgtcgttt 6540
tgccaggtga ccacgttggt caagaaatca cagccgaagc cattaaggtt cttaaagcta 6600
tttctgatgt tcgttccaat gtcaagttcg atttcgaaaa tcatttaatt ggtggtgctg 6660
ctatcgatgc tacaggtgtc ccacttccag atgaggcgct ggaagcctcc aagaaggttg 6720
atgccgtttt gttaggtgct gtgggtggtc ctaaatgggg taccggtagt gttagacctg 6780
aacaaggttt actaaaaatc cgtaaagaac ttcaattgta cgccaactta agaccatgta 6840
actttgcatc cgactctctt ttagacttat ctccaatcaa gccacaattt gctaaaggta 6900
ctgacttcgt tgttgtcaga gaattagtgg gaggtattta ctttggtaag agaaaggaag 6960
acgatggtga tggtgtcgct tgggatagtg aacaatacac cgttccagaa gtgcaaagaa 7020
tcacaagaat ggccgctttc atggccctac aacatgagcc accattgcct atttggtcct 7080
tggataaagc taatgttttg gcctcttcaa gattatggag aaaaactgtg gaggaaacca 7140
tcaagaacga attccctaca ttgaaggttc aacatcaatt gattgattct gccgccatga 7200
tcctagttaa gaacccaacc cacctaaatg gtattataat caccagcaac atgtttggtg 7260
atatcatctc cgatgaagcc tccgttatcc caggttcctt gggtttgttg ccatctgcgt 7320
ccttggcctc tttgccagac aagaacaccg catttggttt gtacgaacca tgccacggtt 7380
ctgctccaga tttgccaaag aataaggtca accctatcgc cactatcttg tctgctgcaa 7440
tgatgttgaa attgtcattg aacttgcctg aagaaggtaa ggccattgaa gatgcagtta 7500
aaaaggtttt ggatgcaggt atcagaactg gtgatttagg tggttccaac agtaccacgg 7560
aagtcggtga tgctgtcgcc gaagaagtta agaaaatcct tgcttaaaaa gattctcttt 7620
ttttatgata tttgtacata aactttataa atgaaattca taatagaaac gacacgaaat 7680
tacaaaatgg aatatgttca tagggtagac gaaactatat acgcaatcta catacattta 7740
tcaagaagga gaaaaaggag gatgtaaagg aatacaggta agcaaattga tactaatggc 7800
tcaacgtgat aaggaaaaag aattgcactt taacattaat attgacaagg aggagggcac 7860
cacacaaaaa gttaggtgta acagaaaatc atgaaactat gattcctaat ttatatattg 7920
gaggattttc tctaaaaaaa aaaaaataca acaaataaaa aacactcaat gacctgacca 7980
tttgatggag tttaagtcaa taccttcttg aaccatttcc cataatggtg aaagttccct 8040
caagaatttt actctgtcag aaacggcctt aacgacgtag tcgacctcct cttcagtact 8100
aaatctacca ataccaaatc tgatggaaga atgggctaat gcatcatcct tacccagcgc 8160
atgtaaaaca taagaaggtt ctagggaagc agatgtacag gctgaacccg aggataatgc 8220
gatatccctt agtgccatca ataaagattc tccttccacg taggcgaaag aaacgttaac 8280
acaccctgga taacgatgat ctggagatcc gttcaacgtg gtatgttcag cggataatag 8340

accttgact aatttatcgg atagtctttt gatgtgagct tggtcgttgt caaattcttt 8400
cttcatcaat ctcgcagctt caccaaatcc cgctaccaat gggggggcca aagtaccaga 8460
tctgctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc 8520
ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc 8580
agcatcgatg aattccacgg actatagact atactagtat actccgtcta ctgtacgata 8640
cacttccgct caggtccttg tcctttaacg aggccttacc actcttttgt tactctattg 8700
atccagctca gcaaaggcag tgtgatctaa gattctatct tcgcgatgta gtaaaactag 8760
ctagaccgag aaagagacta gaaatgcaaa aggcacttct acaatggctg ccatcattat 8820
tatccgatgt gacgctgcag cttctcaatg atattcgaat acgctttgag gagatacagc 8880
ctaatatccg acaaactgtt ttacagattt acgatcgtac ttgttaccca tcattgaatt 8940
ttgaacatcc gaacctggga gttttccctg aaacagatag tatatttgaa cctgtataat 9000
aatatatagt ctagcgcttt acggaagaca atgtatgtat ttcggttcct ggagaaacta 9060
ttgcatctat tgcataggta atcttgcacg tcgcatcccc ggttcatttt ctgcgtttcc 9120
atcttgcact tcaatagcat atctttgtta acgaagcatc tgtgcttcat tttgtagaac 9180
aaaaatgcaa cgcgagagcg ctaatttttc aaacaaagaa tctgagctgc atttttacag 9240
aacagaaatg caacgcgaaa gcgctatttt accaacgaag aatctgtgct tcatttttgt 9300
aaaacaaaaa tgcaacgcga gagcgctaat ttttcaaaca aagaatctga gctgcatttt 9360
tacagaacag aaatgcaacg cgagagcgct attttaccaa caaagaatct atacttcttt 9420
tttgttctac aaaaatgcat cccgagagcg ctatttttct aacaaagcat cttagattac 9480
ttttttttctc ctttgtgcgc tctataatgc agtctcttga taacttttg cactgtaggt 9540
ccgttaaggt tagaagaagg ctactttggt gtctattttc tcttccataa aaaaagcctg 9600
actccacttc ccgcgtttac tgattactag cgaagctgcg ggtgcatttt ttcaagataa 9660
aggcatcccc gattatattc tataccgatg tggattgcgc atactttgtg aacagaaagt 9720
gatagcgttg atgattcttc attggtcaga aaattatgaa cggtttcttc tattttgtct 9780
ctatatacta cgtataggaa atgtttacat tttcgtattg ttttcgattc actctatgaa 9840
tagttcttac tacaattttt ttgtctaaag agtaatacta gagataaaca taaaaaatgt 9900
agaggtcgag tttagatgca agttcaagga gcgaaaggtg gatgggtagg ttatataggg 9960
atatagcaca gagatatata gcaaagagat actttgagc aatgtttgtg gaagcggtat 10020
tcgcaatatt ttagtagctc gttacagtcc ggtgcgtttt tggtttttg aaagtgcgtc 10080
ttcagagcgc ttttggtttt caaaagcgct ctgaagttcc tatactttct agagaatagg 10140
aacttcggaa taggaacttc aaagcgtttc cgaaaacgag cgcttccgaa aatgcaacgc 10200
gagctgcgca catacagctc actgttcacg tcgcacctat atctgcgtgt tgcctgtata 10260
tatatataca tgagaagaac ggcatagtgc gtgtttatgc ttaaatgcgt acttatatgc 10320
gtctatttat gtaggatgaa aggtagtcta gtacctcctg tgatattatc ccattccatg 10380
cggggtatcg tatgcttcct tcagcactac cctttagctg ttctatatgc tgccactcct 10440
caattggatt agtctcatcc ttcaatgcta tcatttcctt tgatattgga tcatatgcat 10500
agtaccgaga aactagtgcg aagtagtgat caggtattgc tgttatctga tgagtatacg 10560
ttgtcctggc cacggcagaa gcacgcttat cgctccaatt tcccacaaca ttagtcaact 10620
ccgttaggcc cttcattgaa agaaatgagg tcatcaaatg tcttccaatg tgagattttg 10680
ggccattttt tatagcaaag attgaataag gcgcattttt cttcaaagct ttattgtacg 10740
atctgactaa gttatctttt aataattggt attcctgttt attgcttgaa gaattgccgg 10800
tcctatttac tcgttttagg actggttcag aattcatcga tgctcactca aaggtcggta 10860
atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag 10920
caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt tt 10962

<210> 234

<211> 3509
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS9008

<400> 234
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggca aacaccggtg gatcctccgg agtggatcta 480
cgcacgctcg gctacagcca gcagcaacag gagaagatca aaccgaaggt tcgttcgaca 540
gtggcgcagc accacgaggc actggtcggc cacgggttta cacacgcgca catcgttgcg 600
ttaagccaac acccggcagc gttagggacc gtcgctgtca agtatcagga catgatcgca 660
gcgttgccag aggcgacaca cgaagcgatc gttggcgtcg gcaaacagtg gtccggcgca 720

cgcgctctgg aggccttgct cacggtggcg ggagagttga gaggtccacc gttacagttg 780
gacacaggcc aacttctcaa gattgcaaaa cgtggcggcg tgaccgcagt ggaggcagtg 840
catgcatggc gcaatgcact gacgggtgcc ccgctcaact tgaccggaga cgcccggggg 900
atcaggtcac gtgcgtctcg gagcattgtt gcccagttat ctcgccctga tccgtcggcc 960
gactgataag tcgaccacca ccaccaccac cactgataag agctcctcga gcctgcagca 1020
gctgaagctt tggacttctg gcgcgcccaa ttcgccctat agtgagtcgt attacgtcgc 1080
gctcactggc cgtcgtttta caacgtcgtg actgggaaaa ccctggcgtt acccaactta 1140
atcgccttgc agcacatccc cctttcgcca gctggcgtaa tagcgaagag gcccgcaccg 1200
aaacgccctt cccaacagtt gcgcagcctg aatggcgaat gggagcgccc tgtagcggcg 1260
cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgccc 1320
tagcgcccgc tcctttcgct ttcttccctt cctttctcgc cacgttcgcc ggctttcccc 1380
gtcaagctct aaatcggggg ctccctttag ggttccgatt tagtgcttta cggcacctcg 1440
accccaaaaa acttgattag ggtgatggtt ggcctgtagt gggccatagc cctgatagac 1500
ggtttttcgc cctttgacgt tggagtccac gttctttaat agtggactct tgttccaaac 1560
tggaacaaca ctcaacccta tctcggtcta ttcttttgat ttataaggga ttttgccgat 1620
ttcggcctat tggttaaaaa atgagctgat ttaacaaaaa tttaacgcga attttaacaa 1680
aatattaacg cttacaattt aggtggcact tttcggggaa atgtgcgcgg aacccctatt 1740
tgtttatttt tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa 1800
atgcttcaat aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt 1860
attccctttt ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa 1920
gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac 1980
agcggtaaga tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt 2040
aaagttctgc tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggt 2100
cgccgcatac actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat 2160
cttacggatg gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac 2220
actgcggcca acttacttct gacaacgatc ggaggaccga aggagctaac cgcttttttg 2280
cacaacatgg gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc 2340
ataccaaacg acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa 2400
ctattaactg gcgaactact tactctagct tcccggcaac aattaataga ctggatggag 2460
gcggataaag ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct 2520
gataaatctg gagccggtga gcgtggttct cgcggtatca ttgcagcact ggggccagat 2580
ggtaagccct cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa 2640
cgaaatagac agatcgctga gataggtgcc tcactgatta agcattggta actgtcagac 2700
caagtttact catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc 2760
taggtgaaga tccttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc 2820
cactgagcgt cagaccccgt agaaaagatc aaaggatctt cttgagatcc ttttttctg 2880
cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg 2940
gatcaagagc taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca 3000
aatactgttc ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg 3060
cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg 3120
tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga 3180
acgggggttt cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac 3240
ctacagcgtg agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat 3300
ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc 3360
tggtatcttt atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga 3420
tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc 3480
ctggcctttt gctggccttt tgctcacat 3509

<210> 235

<211> 4049
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS12730

<400> 235
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggcc aagtctggca tctaccagat taagaacacc 480
ctgaataaca aagtgtatgt tggatctgcc aaggactttg agaaaagatg gaagaggcac 540

ttcaaggatc tggaaaaggg ttgccacagc tccatcaagc tccaaaggtc tttcaacaaa 600
catggcaatg tctttgagtg ctccattttg gaagagatcc cttatgaaaa ggacctcatt 660
atagagaggg aaaacttctg gatcaaggag ctgaactcaa aaatcaatgg atacaacatt 720
gcagatgcca catttggtga cacctgttct actcaccccc ttaaggaaga gatcattaaa 780
aagaggtctg agactgtgaa ggctaagatg ttgaaactgg ggcctgatgg cagaaaggca 840
ctctactcta agccaggtag caaaaatgga aggtggaacc ctgagaccca caagttctgc 900
aaatgtgggg tgagaataca gacatcagcc tacacttgca gcaagtgtag gaacagatca 960
ggagagaata actccttttt caaccacaag cattctcaag gtccatccgg agtggatcta 1020
cgcacgctcg gctacagcca gcagcaacag gagaagatca aaccgaaggt tcgttcgaca 1080
gtggcgcagc accacgaggc actggtcggc cacgggttta cacacgcgca catcgttgcg 1140
ttaagccaac acccggcagc gttagggacc gtcgctgtca agtatcagga catgatcgca 1200
gcgttgccag aggcgacaca cgaagcgatc gttggcgtcg gcaaacagtg gtccggcgca 1260
cgcgctctgg aggccttgct cacggtggcg ggagagttga gaggtccacc gttacagttg 1320
gacacaggcc aacttctcaa gattgcaaaa cgtggcggcg tgaccgcagt ggaggcagtg 1380
catgcatggc gcaatgcact gacgggtgcc ccgctcaact tgaccggaga cgcccggggg 1440
atcaggtcac gtgcgtctcg gagcattgtt gcccagttat ctcgccctga tccgtcggcc 1500
gactgataag tcgaccacca ccaccaccac cactgataag agctcctcga gcctgcagca 1560
gctgaagctt tggacttctg gcgcgcccaa ttcgccctat agtgagtcgt attacgtcgc 1620
gctcactggc cgtcgtttta caacgtcgtg actgggaaaa ccctggcgtt acccaactta 1680
atcgccttgc agcacatccc cctttcgcca gctggcgtaa tagcgaagag gcccgcaccg 1740
aaacgccctt cccaacagtt gcgcagcctg aatggcgaat gggagcgccc tgtagcggcg 1800
cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgccc 1860
tagcgcccgc tcctttcgct ttcttccctt cctttctcgc cacgttcgcc ggctttcccc 1920
gtcaagctct aaatcggggg ctccctttag ggttccgatt tagtgcttta cggcacctcg 1980
accccaaaaa acttgattag ggtgatggtt ggcctgtagt gggccatagc cctgatagac 2040
ggtttttcgc cctttgacgt tggagtccac gttctttaat agtggactct tgttccaaac 2100
tggaacaaca ctcaacccta tctcggtcta ttcttttgat ttataaggga ttttgccgat 2160
ttcggcctat tggttaaaaa atgagctgat ttaacaaaaa tttaacgcga attttaacaa 2220
aatattaacg cttacaattt aggtggcact tttcggggaa atgtgcgcgg aacccctatt 2280
tgtttatttt tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa 2340
atgcttcaat aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt 2400
attccctttt ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa 2460
gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac 2520
agcggtaaga tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt 2580
aaagttctgc tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggt 2640
cgccgcatac actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat 2700
cttacggatg gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac 2760
actgcggcca acttacttct gacaacgatc ggaggaccga aggagctaac cgcttttttg 2820
cacaacatgg gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc 2880
ataccaaacg acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa 2940
ctattaactg gcgaactact tactctagct tcccggcaac aattaataga ctggatggag 3000
gcggataaag ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct 3060
gataaatctg gagccggtga gcgtggttct cgcggtatca ttgcagcact ggggccagat 3120
ggtaagccct cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa 3180
cgaaatagac agatcgctga gataggtgcc tcactgatta agcattggta actgtcagac 3240
caagtttact catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc 3300
taggtgaaga tccttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc 3360
cactgagcgt cagaccccgt agaaaagatc aaaggatctt cttgagatcc ttttttttctg 3420
cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg 3480
gatcaagagc taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca 3540
aatactgttc ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg 3600
cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg 3660
tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga 3720
acgggggttt cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac 3780
ctacagcgtg agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat 3840
ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc 3900
tggtatcttt atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga 3960
tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc 4020
ctggccttttt gctggccttt tgctcacat 4049

<210> 236

<211> 5801
<212> DNA
<213> Artificial sequence

<220>

<223> pCLS12731

<400> 236

ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggcc aagtctggca tctaccagat taagaacacc 480
ctgaataaca aagtgtatgt tggatctgcc aaggactttg agaaaagatg gaagaggcac 540
ttcaaggatc tggaaaaggg ttgccacagc tccatcaagc tccaaaggtc tttcaacaaa 600
catggcaatg tctttgagtg ctccattttg gaagagatcc cttatgaaaa ggacctcatt 660
atagagaggg aaaacttctg gatcaaggag ctgaactcaa aaatcaatgg atacaacatt 720
gcagatgcca catttggtga cacctgttct actcaccccc ttaaggaaga gatcattaaa 780
aagaggtctg agactgtgaa ggctaagatg ttgaaactgg ggcctgatgg cagaaaggca 840
ctctactcta agccaggtag caaaaatgga aggtggaacc ctgagaccca caagttctgc 900
aaatgtgggg tgagaataca gacatcagcc tacacttgca gcaagtgtag gaacagatca 960
ggagagaata actccttttt caaccacaag cattctcaag gtccatccgg agtggatcta 1020
cgcacgctcg gctacagcca gcagcaacag gagaagatca aaccgaaggt tcgttcgaca 1080
gtggcgcagc accacgaggc actggtcggc cacgggttta cacacgcgca catcgttgcg 1140
ttaagccaac acccggcagc gttagggacc gtcgctgtca agtatcagga catgatcgca 1200
gcgttgccag aggcgacaca cgaagcgatc gttggcgtcg gcaaacagtg gtccggcgca 1260
cgcgctctgg aggccttgct cacggtggcg ggagagttga gaggtccacc gttacagttg 1320
gacacaggcc aacttctcaa gattgcaaaa cgtggcggcg tgaccgcagt ggaggcagtg 1380
catgcatggc gcaatgcact gacgggtgcc ccgctcaact tgaccccgga gcaggtggtg 1440
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 1500
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 1560
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 1620
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg 1680
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 1740
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 1800
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 1860
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 1920
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag 1980
gcgctggaga cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 2040
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 2100
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 2160
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 2220
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 2280
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 2340
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 2400
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 2460
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 2520
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagcaatatt 2580
ggtggcaagc aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac 2640
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg 2700
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag 2760
gtggtggcca tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg 2820
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 2880
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2940
gcccacggct tgaccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 3000
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 3060
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 3120
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc ctcagcaggt ggtggccatc 3180
gccagcaatg gcggcggcag gccggcgctg gagagcattg ttgcccagtt atctcgccct 3240
gatccgtcgg ccgactgata agtcgaccac caccaccacc accactgata agagctcctc 3300
gagcctgcag cagctgaagc tttggacttc tggcgcgccc aattcgccct atagtgagtc 3360
gtattacgtc gcgctcactg gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg 3420
ttacccaact taatcgcctt gcagcacatc ccccttcgc cagctggcgt aatagcgaag 3480
aggcccgcac cgaaacgccc ttcccaacag ttgcgcagcc tgaatggcga atgggagcgc 3540
cctgtagcgg cgcattaagc gcggcgggtg tggtggttac gcgcagcgtg accgctacac 3600
ttgccagcgc cctagcgccc gctcctttcg ctttcttccc ttcctttctc gccacgttcg 3660
ccggctttcc ccgtcaagct ctaaatcggg ggctcccttt agggttccga tttagtgctt 3720
tacggcacct cgaccccaaa aaacttgatt agggtgatgg ttggcctgta gtgggccata 3780
gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta atagtggact 3840
cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg 3900

gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc 3960
gaattttaac aaaatattaa cgcttacaat ttaggtggca cttttcgggg aaatgtgcgc 4020
ggaaccccta tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa 4080
taaccctgat aaatgcttca ataatattga aaaaggaaga gtatgagtat tcaacatttc 4140
cgtgtcgccc ttattccctt ttttgcggca ttttgccttc ctgtttttgc tcacccagaa 4200
acgctggtga aagtaaaaga tgctgaagat cagttgggtg cacgagtggg ttacatcgaa 4260
ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg 4320
atgagcactt ttaaagttct gctatgtggc gcggtattat cccgtattga cgccgggcaa 4380
gagcaactcg gtcgccgcat acactattct cagaatgact tggttgagta ctcaccagtc 4440
acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc 4500
atgagtgata acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta 4560
accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag 4620
ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgcctgtagc aatggcaaca 4680
acgttgcgca aactattaac tggcgaacta cttactctag cttcccggca acaattaata 4740
gactggatgg aggcggataa agttgcagga ccacttctgc gctcggccct tccggctggc 4800
tggtttattg ctgataaatc tggagccggt gagcgtggtt ctcgcggtat cattgcagca 4860
ctggggccag atggtaagcc ctcccgtatc gtagttatct acacgacggg gagtcaggca 4920
actatggatg aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg 4980
taactgtcag accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa 5040
tttaaaagga tctaggtgaa gatccttttt gataatctca tgaccaaaat cccttaacgt 5100
gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat 5160
ccttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg 5220
gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga 5280
gcgcagatac caaatactgt tcttctagtg tagccgtagt taggccacca cttcaagaac 5340
tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt 5400
ggcgataagt cgtgtcttac cgggttggac tcaagacgat agttaccgga taaggcgcag 5460
cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc 5520
gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaag 5580
gcggacaggt atccggtaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca 5640
gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt 5700
cgatttttgt gatgctcgtc agggggggcgg agcctatgga aaaacgccag caacgcggcc 5760
tttttacggt tcctggcctt ttgctggcct tttgctcaca t 5801

<210> 237

<211> 11112
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS8522

<400> 237
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
tttttgtttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440

gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg ccaagtctgg catctaccag 2460
attaagaaca ccctgaataa caaagtgtat gttggatctg ccaaggactt tgagaaaaga 2520
tggaagaggc acttcaagga tctggaaaag ggttgccaca gctccatcaa gctccaaagg 2580
tctttcaaca aacatggcaa tgtctttgag tgctccattt tggaagagat cccttatgaa 2640
aaggacctca ttatagagag ggaaaacttc tggatcaagg agctgaactc aaaaatcaat 2700
ggatacaaca ttgcagatgc cacatttggt gacacctgtt ctactcaccc ccttaaggaa 2760
gagatcatta aaaagaggtc tgagactgtg aaggctaaga tgttgaaact ggggcctgat 2820
ggcagaaagg cactctactc taagccaggt agcaaaaatg gaaggtggaa ccctgagacc 2880
cacaagttct gcaaatgtgg ggtgagaata cagacatcag cctacacttg cagcaagtgt 2940
aggaacagat caggagagaa taactccttt ttcaaccaca agcattctca aggtccatcc 3000
ggagtggatc tacgcacgct cggctacagc cagcagcaac aggagaagat caaaccgaag 3060
gttcgttcga cagtggcgca gcaccacgag gcactggtcg gccacgggtt tacacacgcg 3120
cacatcgttg cgttaagcca acacccggca gcgttaggga ccgtcgctgt caagtatcag 3180
gacatgatcg cagcgttgcc agaggcgaca cacgaagcga tcgttggcgt cggcaaacag 3240
tggtccggcg cacgcgctct ggaggccttg ctcacggtgg cgggagagtt gagaggtcca 3300
ccgttacagt tggacacagg ccaacttctc aagattgcaa aacgtggcgg cgtgaccgca 3360
gtggaggcag tgcatgcatg gcgcaatgca ctgacgggtg ccccgctcaa cttgaccccg 3420
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 3480
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 3540
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3600
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 3660
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 3720
acccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 3780
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 3840
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 3900
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagcaatatt 3960
ggtggcaagc aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac 4020
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg 4080
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag 4140
gtggtggcca tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg 4200
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 4260
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 4320
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 4380
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 4440
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 4500
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 4560
gccagcaata ttggtggcaa gcaggcgctg gagacggtgc aggcgctgtt gccggtgctg 4620
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 4680
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 4740
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 4800
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 4860
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 4920
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 4980
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 5040
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg 5100
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac ccctcagcag 5160
gtggtggcca tcgccagcaa tggcggcggc aggccggcgc tggagagcat tgttgcccag 5220
ttatctcgcc ctgatccgtc ggccgactga taactcgagc gatcctctag acgagctcct 5280
cgagcctgca gcagctgaag ctttggactt cttcgccaga ggtttggtca agtctccaat 5340
caaggttgtc ggcttgtcta ccttgccaga aatttacgaa aagatggaaa agggtcaaat 5400
cgttggtaga tacgttgttg acacttctaa ataagcgaat ttcttatgat ttatgatttt 5460
tattattaaa taagttataa aaaaaataag tgtatacaaa ttttaaagtg actcttaggt 5520

tttaaaacga aaattcttat tcttgagtaa ctctttcctg taggtcaggt tgctttctca 5580
ggtatagcat gaggtcgctc ttattgacca cacctctacc ggcatgcaag cttggcgtaa 5640
tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata 5700
cgagccggaa gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta 5760
attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gcagatctat 5820
tacattatgg gtggtatgtt ggaataaaaa tcaactatca tctactaact agtatttacg 5880
ttactagtat attatcatat acggtgttag aagatgacgc aaatgatgag aaatagtcat 5940
ctaaattagt ggaagctgaa acgcaaggat tgataatgta ataggatcaa tgaatattaa 6000
catataaaat gatgataata atatttatag aattgtgtag aattgcagat tccctttttat 6060
ggattcctaa atcctcgagg agaacttcta gtatatctac atacctaata ttattgcctt 6120
attaaaaatg gaatcccaac aattacatca aaatccacat tctcttcaaa atcaattgtc 6180
ctgtacttcc ttgttcatgt gtgttcaaaa acgttatatt tataggataa ttatactcta 6240
tttctcaaca agtaattggt tgtttggccg agcggtctaa ggcgcctgat tcaagaaata 6300
tcttgaccgc agttaactgt gggaatactc aggtatcgta agatgcaaga gttcgaatct 6360
cttagcaacc attatttttt tcctcaacat aacgagaaca cacaggggcg ctatcgcaca 6420
gaatcaaatt cgatgactgg aaatttttg ttaatttcag aggtcgcctg acgcatatac 6480
ctttttcaac tgaaaaattg ggagaaaaag gaaaggtgag agccgcggaa ccggctttc 6540
atatagaata gagaagcgtt catgactaaa tgcttgcatc acaatacttg aagttgacaa 6600
tattatttaa ggacctattg ttttttccaa taggtggtta gcaatcgtct tactttctaa 6660
cttttcttac cttttacatt tcagcaatat atatatatat atttcaagga tataccattc 6720
taatgtctgc ccctaagaag atcgtcgttt tgccaggtga ccacgttggt caagaaatca 6780
cagccgaagc cattaaggtt cttaaagcta tttctgatgt tcgttccaat gtcaagttcg 6840
atttcgaaaa tcatttaatt ggtggtgctg ctatcgatgc tacaggtgtc ccacttccag 6900
atgaggcgct ggaagcctcc aagaaggttg atgccgtttt gttaggtgct gtgggtggtc 6960
ctaaatgggg taccggtagt gttagacctg aacaaggttt actaaaaatc cgtaaagaac 7020
ttcaattgta cgccaactta agaccatgta actttgcatc cgactctctt ttagacttat 7080
ctccaatcaa gccacaattt gctaaaggta ctgacttcgt tgttgtcaga gaattagtgg 7140
gaggtattta ctttggtaag agaaaggaag acgatggtga tggtgtcgct tgggatagtg 7200
aacaatacac cgttccagaa gtgcaaagaa tcacaagaat ggccgctttc atggccctac 7260
aacatgagcc accattgcct atttggtcct tggataaagc taatgttttg gcctcttcaa 7320
gattatggag aaaaactgtg gaggaaacca tcaagaacga attccctaca ttgaaggttc 7380
aacatcaatt gattgattct gccgccatga tcctagttaa gaacccaacc cacctaaatg 7440
gtattataat caccagcaac atgtttggtg atatcatctc cgatgaagcc tccgttatcc 7500
caggttcctt gggtttgttg ccatctgcgt ccttggcctc tttgccagac aagaacaccg 7560
catttggttt gtacgaacca tgccacggtt ctgctccaga tttgccaaag aataaggtca 7620
accctatcgc cactatcttg tctgctgcaa tgatgttgaa attgtcattg aacttgcctg 7680
aagaaggtaa ggccattgaa gatgcagtta aaaaggtttt ggatgcaggt atcagaactg 7740
gtgatttagg tggttccaac agtaccacgg aagtcggtga tgctgtcgcc gaagaagtta 7800
agaaaatcct tgcttaaaaa gattctcttt ttttatgata tttgtacata aactttataa 7860
atgaaattca taatagaaac gacacgaaat tacaaaatgg aatatgttca tagggtagac 7920
gaaactatat acgcaatcta catacattta tcaagaagga gaaaaaggag gatgtaaagg 7980
aatacaggta agcaaattga tactaatggc tcaacgtgat aaggaaaaag aattgcactt 8040
taacattaat attgacaagg aggagggcac cacacaaaaa gttaggtgta acagaaaatc 8100
atgaaactat gattcctaat ttatatattg gaggattttc tctaaaaaaa aaaaaataca 8160
acaaataaaa aacactcaat gacctgacca tttgatggag tttaagtcaa taccttcttg 8220
aaccatttcc cataatggtg aaagttccct caagaatttt actctgtcag aaacggcctt 8280
aacgacgtag tcgacctcct cttcagtact aaatctacca ataccaaatc tgatggaaga 8340
atgggctaat gcatcatcct tacccagcgc atgtaaaaca taagaaggtt ctagggaagc 8400
agatgtacag gctgaacccg aggataatgc gatatccctt agtgccatca ataaagattc 8460
tccttccacg taggcgaaag aaacgttaac acaccctgga taacgatgat ctggagatcc 8520
gttcaacgtg gtatgttcag cggataatag acctttgact aatttatcgg atagtctttt 8580
gatgtgagct tggtcgttgt caaattcttt cttcatcaat ctcgcagctt caccaaatcc 8640
cgctaccaat gggggggcca aagtaccaga tctgctgcat taatgaatcg gccaacgcgc 8700
ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg 8760
ctcggtcgtt cggctgcggc gagcggtatc agcatcgatg aattccacgg actatagact 8820
atactagtat actccgtcta ctgtacgata cacttccgct caggtccttg tcctttaacg 8880
aggccttacc actcttttgt tactctattg atccagctca gcaaaggcag tgtgatctaa 8940
gattctatct tcgcgatgta gtaaaactag ctagaccgag aaagagacta gaaatgcaaa 9000
aggcacttct acaatggctg ccatcattat tatccgatgt gacgctgcag cttctcaatg 9060
atattcgaat acgctttgag gagatacagc ctaatatccg acaaactgtt ttacagattt 9120
acgatcgtac ttgttaccca tcattgaatt ttgaacatcc gaacctggga gttttccctg 9180
aaacagatag tatatttgaa cctgtataat aatatatagt ctagcgcttt acggaagaca 9240
atgtatgtat ttcggttcct ggagaaacta ttgcatctat tgcataggta atcttgcacg 9300
tcgcatcccc ggttcatttt ctgcgtttcc atcttgcact tcaatagcat atctttgtta 9360
acgaagcatc tgtgcttcat tttgtagaac aaaaatgcaa cgcgagagcg ctaatttttc 9420
aaacaaagaa tctgagctgc atttttacag aacagaaatg caacgcgaaa gcgctatttt 9480
accaacgaag aatctgtgct tcatttttgt aaaacaaaaa tgcaacgcga gagcgctaat 9540
ttttcaaaca aagaatctga gctgcatttt tacagaacag aaatgcaacg cgagagcgct 9600

attttaccaa caaagaatct atacttcttt tttgttctac aaaaatgcat cccgagagcg 9660
ctatttttct aacaaagcat cttagattac tttttttctc ctttgtgcgc tctataatgc 9720
agtctcttga taacttttg cactgtaggt ccgttaaggt tagaagaagg ctactttggt 9780
gtctattttc tcttccataa aaaaagcctg actccacttc ccgcgtttac tgattactag 9840
cgaagctgcg ggtgcatttt ttcaagataa aggcatcccc gattatattc tataccgatg 9900
tggattgcgc atactttgtg aacagaaagt gatagcgttg atgattcttc attggtcaga 9960
aaattatgaa cggtttcttc tattttgtct ctatatacta cgtataggaa atgtttacat 10020
tttcgtattg ttttcgattc actctatgaa tagttcttac tacaattttt ttgtctaaag 10080
agtaatacta gagataaaca taaaaaatgt agaggtcgag tttagatgca agttcaagga 10140
gcgaaaggtg gatgggtagg ttatataggg atatagcaca gagatatata gcaaagagat 10200
acttttgagc aatgtttgtg gaagcggtat tcgcaatatt ttagtagctc gttacagtcc 10260
ggtgcgtttt tggttttttg aaagtgcgtc ttcagagcgc ttttggtttt caaaagcgct 10320
ctgaagttcc tatactttct agagaatagg aacttcggaa taggaacttc aaagcgtttc 10380
cgaaaacgag cgcttccgaa aatgcaacgc gagctgcgca catacagctc actgttcacg 10440
tcgcacctat atctgcgtgt tgcctgtata tatatataca tgagaagaac ggcatagtgc 10500
gtgtttatgc ttaaatgcgt acttatatgc gtctatttat gtaggatgaa aggtagtcta 10560
gtacctcctg tgatattatc ccattccatg cggggtatcg tatgcttcct tcagcactac 10620
cctttagctg ttctatatgc tgccactcct caattggatt agtctcatcc ttcaatgcta 10680
tcatttcctt tgatattgga tcatatgcat agtaccgaga aactagtgcg aagtagtgat 10740
caggtattgc tgttatctga tgagtatacg ttgtcctggc cacggcagaa gcacgcttat 10800
cgctccaatt tcccacaaca ttagtcaact ccgttaggcc cttcattgaa agaaatgagg 10860
tcatcaaatg tcttccaatg tgagattttg ggccattttt tatagcaaag attgaataag 10920
gcgcattttt cttcaaagct ttattgtacg atctgactaa gttatctttt aataattggt 10980
attcctgttt attgcttgaa gaattgccgg tcctatttac tcgttttagg actggttcag 11040
aattcatcga tgctcactca aaggtcggta atacggttat ccacagaatc aggggataac 11100
gcaggaaaga ac 11112

<210> 238

<211> 61
<212> DNA
<213> Artificial sequence

<220>
<223> Avr25RAGT2R

<400> 238
tatataaacc taaccctcta ctagcatgaa ggtaccttgt cgttataagt aaccataaac 60
a 61

<210> 239

<211> 39
<212> DNA
<213> Artificial sequence

<220>
<223> CMP_G001

<400> 239
acacgcaaac acaaatacac agcggccttg ccaccatgg 39

<210> 240

<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> CMP_G068

<400> 240
ggtatttccg gatggacctt gagaatgctt gtggttg 37

<210> 241

<211> 6017
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS7763

<400> 241
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca 780
tttccccgaa aagtgccacc tgacgtccga tcaaaaatca tcgcttcgct gattaattac 840
cccagaaata aggctaaaaa actaatcgca ttatcatcct atggttgtta atttgattcg 900
ttcatttgaa ggtttgtggg gccaggttac tgccaatttt tcctcttcat aaccataaaa 960
gctagtattg tagaatcttt attgttcgga gcagtgcggc gcgaggcaca tctgcgtttc 1020
aggaacgcga ccggtgaaga cgaggacgca cggaggagag tcttccttcg gagggctgtc 1080
acccgctcgg cggcttctaa tccgtacttc aatatagcaa tgagcagtta agcgtattac 1140
tgaaagttcc aaagagaagg ttttttttagg ctaatcgacc tcgagcagat ccgccaggcg 1200
tgtatatagc gtggatggcc aggcaacttt agtgctgaca catacaggca tatatatatg 1260
tgtgcgacga cacatgatca tatggcatgc atgtgctctg tatgtatata aaactcttgt 1320
tttcttcttt tctctaaata ttctttcctt atacattagg tcctttgtag cataaattac 1380
tatacttcta tagacacgca aacacaaata cacagcggcc ttgccaccat ggccggcgcg 1440
cccactatgt gtgcggccga ctgataactc gagcgatcct ctagacgagc tcctcgagcc 1500
tgcagcagct gaagctttgg acttcttcgc cagaggtttg gtcaagtctc caatcaaggt 1560
tgtcggcttg tctaccttgc cagaaattta cgaaaagatg gaaaagggtc aaatcgttgg 1620
tagatacgtt gttgacactt ctaaataagc gaatttctta tgatttatga tttttattat 1680
taaataagtt ataaaaaaaa taagtgtata caaattttaa agtgactctt aggttttaaa 1740
acgaaaattc ttattcttga gtaactcttt cctgtaggtc aggttgcttt ctcaggtata 1800
gcatgaggtc gctcttattg accacacctc taccggcatg caagcttggc gtaatcatgg 1860
tcatagctgt ttcctgtgtg aaattgttat ccgctcacaa ttccacacaa catacgagcc 1920
ggaagcataa agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaattgcg 1980
ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt gccagcagat ctgtttagct 2040
tgcctcgtcc ccgccgggtc acccggccag cgacatggag gcccagaata ccctccttga 2100
cagtcttgac gtgcgcagct caggggcatg atgtgactgt cgcccgtaca tttagcccat 2160
acatccccat gtataatcat ttgcatccat acattttgat ggccgcacgg cgcgaagcaa 2220
aaattacggc tcctcgctgc agacctgcga gcagggaaac gctcccctca cagacgcgtt 2280
gaattgtccc cacgccgcgc ccctgtagag aaatataaaa ggttaggatt tgccactgag 2340
gttcttcttt catatacttc cttttaaaat cttgctagga tacagttctc acatcacatc 2400
cgaacataaa caaccatgca tgggtaagga aaagactcac gtttcgaggc cgcgattaaa 2460
ttccaacatg gatgctgatt tatatgggta taaatgggct cgcgataatg tcgggcaatc 2520
aggtgcgaca atctatcgat tgtatgggaa gcccgatgcg ccagagttgt ttctgaaaca 2580
tggcaaaggt agcgttgcca atgatgttac agatgagatg gtcagactaa actggctgac 2640
ggaatttatg cctcttccga ccatcaagca ttttatccgt actcctgatg atgcatggtt 2700
actcaccact gcgatccccg gcaaaacagc attccaggta ttagaagaat atcctgattc 2760
aggtgaaaat attgttgatg cgctggcagt gttcctgcgc cggttgcatt cgattcctgt 2820
ttgtaattgt ccttttaaca gcgatcgcgt atttcgcctc gctcaggcgc aatcacgaat 2880
gaataacggt ttggttgatg cgagtgattt tgatgacgag cgtaatggct ggcctgttga 2940
acaagtctgg aaagaaatgc ataagctttt gccattctca ccggattcag tcgtcactca 3000
tggtgatttc tcacttgata accttatttt tgacgagggg aaattaatag gttgtattga 3060
tgttggacga gtcggaatcg cagaccgata ccaggatctt gccatcctat ggaactgcct 3120
cggtgagttt tctccttcat tacagaaacg gctttttcaa aaatatggta ttgataatcc 3180
tgatatgaat aaattgcagt ttcatttgat gctcgatgag tttttctaat cagtactgac 3240
aataaaaaga ttcttgtttt caagaacttg tcatttgtat agttttttta tattgtagtt 3300
gttctatttt aatcaaatgt tagcgtgatt tatatttttt ttcgcctcga catcatctgc 3360
ccagatgcga agttaagtgc gcagaaagta atatcatgcg tcaatcgtat gtgaatgctg 3420
gtcgctatac tgctgtcgat tcgatactaa cgccgccatc cagtgtcgaa aacgagctcg 3480

aattcatcga tgatatcaga tccactagtg gcctatgcga ccgcggatct gccggtctcc 3540
ctatagtgag tcgtattaat ttcgataagc caggttaacc tgcattaatg aatcggccaa 3600
cgcgcgggga gaggcggttt gcgtattggg cgctcttccg cttcctcgct cactgactcg 3660
ctgcgctcgg tcgttcggct gcggcgagcg gtatcagcat cgatgaattc cacggactat 3720
agactatact agtatactcc gtctactgta cgatacactt ccgctcaggt ccttgtcctt 3780
taacgaggcc ttaccactct tttgttactc tattgatcca gctcagcaaa ggcagtgtga 3840
tctaagattc tatcttcgcg atgtagtaaa actagctaga ccgagaaaga gactagaaat 3900
gcaaaaggca cttctacaat ggctgccatc attattatcc gatgtgacgc tgcagcttct 3960
caatgatatt cgaatacgct ttgaggagat acagcctaat atccgacaaa ctgttttaca 4020
gatttacgat cgtacttgtt acccatcatt gaattttgaa catccgaacc tgggagtttt 4080
ccctgaaaca gatagtatat ttgaacctgt ataataatat atagtctagc gctttacgga 4140
agacaatgta tgtatttcgg ttcctggaga aactattgca tctattgcat aggtaatctt 4200
gcacgtcgca tccccggttc attttctgcg tttccatctt gcacttcaat agcatatctt 4260
tgttaacgaa gcatctgtgc ttcattttgt agaacaaaaa tgcaacgcga gagcgctaat 4320
ttttcaaaca aagaatctga gctgcatttt tacagaacag aaatgcaacg cgaaagcgct 4380
attttaccaa cgaagaatct gtgcttcatt tttgtaaaac aaaaatgcaa cgcgagagcg 4440
ctaatttttc aaacaaagaa tctgagctgc atttttacag aacagaaatg caacgcgaga 4500
gcgctatttt accaacaaag aatctatact tctttttttgt tctacaaaaa tgcatcccga 4560
gagcgctatt tttctaacaa agcatcttag attacttttt ttctcctttg tgcgctctat 4620
aatgcagtct cttgataact ttttgcactg taggtccgtt aaggttagaa gaaggctact 4680
ttggtgtcta ttttctcttc cataaaaaaa gcctgactcc acttcccgcg tttactgatt 4740
actagcgaag ctgcgggtgc atttttttcaa gataaaggca tccccgatta tattctatac 4800
cgatgtggat tgcgcatact ttgtgaacag aaagtgatag cgttgatgat tcttcattgg 4860
tcagaaaatt atgaacggtt tcttctattt tgtctctata tactacgtat aggaaatgtt 4920
tacattttcg tattgttttc gattcactct atgaatagtt cttactacaa tttttttgtc 4980
taaagagtaa tactagagat aaacataaaa aatgtagagg tcgagtttag atgcaagttc 5040
aaggagcgaa aggtggatgg gtaggttata tagggatata gcacagagat atatagcaaa 5100
gagatacttt tgagcaatgt ttgtggaagc ggtattcgca atattttagt agctcgttac 5160
agtccggtgc gtttttggtt ttttgaaagt gcgtcttcag agcgcttttg gttttcaaaa 5220
gcgctctgaa gttcctatac tttctagaga ataggaactt cggaatagga acttcaaagc 5280
gtttccgaaa acgagcgctt ccgaaaatgc aacgcgagct gcgcacatac agctcactgt 5340
tcacgtcgca cctatatctg cgtgttgcct gtatatatat atacatgaga agaacggcat 5400
agtgcgtgtt tatgcttaaa tgcgtactta tatgcgtcta tttatgtagg atgaaaggta 5460
gtctagtacc tcctgtgata ttatcccatt ccatgcgggg tatcgtatgc ttccttcagc 5520
actacccttt agctgttcta tatgctgcca ctcctcaatt ggattagtct catccttcaa 5580
tgctatcatt tcctttgata ttggatcata tgcatagtac cgagaaacta gtgcgaagta 5640
gtgatcaggt attgctgtta tctgatgagt atacgttgtc ctggccacgg cagaagcacg 5700
cttatcgctc caatttccca caacattagt caactccgtt aggcccttca ttgaaagaaa 5760
tgaggtcatc aaatgtcttc caatgtgaga ttttgggcca ttttttatag caaagattga 5820
ataaggcgca tttttcttca aagctttatt gtacgatctg actaagttat cttttaataa 5880
ttggtattcc tgtttattgc ttgaagaatt gccggtccta tttactcgtt ttaggactgg 5940
ttcagaattc atcgatgctc actcaaaggt cggtaatacg gttatccaca gaatcagggg 6000
ataacgcagg aaagaac 6017

<210> 242

<211> 6262
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS9922

<400> 242
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
tttttgtttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca 780

tttccccgaa aagtgccacc tgacgtccga tcaaaaatca tcgcttcgct gattaattac 840
cccagaaata aggctaaaaa actaatcgca ttatcatcct atggttgtta atttgattcg 900
ttcatttgaa ggtttgtggg gccaggttac tgccaatttt tcctcttcat aaccataaaa 960
gctagtattg tagaatcttt attgttcgga gcagtgcggc gcgaggcaca tctgcgtttc 1020
aggaacgcga ccggtgaaga cgaggacgca cggaggagag tcttccttcg gagggctgtc 1080
acccgctcgg cggcttctaa tccgtacttc aatatagcaa tgagcagtta agcgtattac 1140
tgaaagttcc aaagagaagg ttttttagg ctaatcgacc tcgagcagat ccgccaggcg 1200
tgtatatagc gtggatggcc aggcaacttt agtgctgaca catacaggca tatatatatg 1260
tgtgcgacga cacatgatca tatggcatgc atgtgctctg tatgtatata aaactcttgt 1320
tttcttcttt tctctaaata ttctttcctt atacattagg tcctttgtag cataaattac 1380
tatacttcta tagacacgca aacacaaata cacagcggcc ttgccaccat gggatccgaa 1440
ctgaagaact ccatcagcga ttacaccgaa gccgagttcg tgcagctgtt gaaagaaatt 1500
gagaaagaga acgttgctgc caccgatgac gtgctggatg ttttgctgga acattttgtg 1560
aagatcactg aacatccaga tggcactgac ctgatctact acccgagcga taaccgtgat 1620
gacagcccag aaggcattgt taaagagatc aaagaatgga gagctgccaa cggcaagcca 1680
ggcttcaaac agggttgacg gccgactgat aactcgagcg atcctctaga cgagctcctc 1740
gagcctgcag cagctgaagc tttggacttc ttcgccagag gtttggtcaa gtctccaatc 1800
aaggttgtcg gcttgtctac cttgccagaa atttacgaaa agatggaaaa gggtcaaatc 1860
gttggtagat acgttgttga cacttctaaa taagcgaatt tcttatgatt tatgattttt 1920
attattaaat aagttataaa aaaaataagt gtatacaaat tttaaagtga ctcttaggtt 1980
ttaaaacgaa aattcttatt cttgagtaac tctttcctgt aggtcaggtt gctttctcag 2040
gtatagcatg aggtcgctct tattgaccac acctctaccg gcatgcaagc ttggcgtaat 2100
catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac 2160
gagccggaag cataaagtgt aaagcctggg gtgcctaatg agtgagctaa ctcacattaa 2220
ttgcgttgcg ctcactgccc gctttccagt cgggaaacct gtcgtgccag cagatctgtt 2280
tagcttgcct cgtccccgcc gggtcacccg gccagcgaca tggaggccca gaataccctc 2340
cttgacagtc ttgacgtgcg cagctcaggg gcatgatgtg actgtcgccc gtacatttag 2400
cccatacatc cccatgtata atcatttgca tccatacatt ttgatggccg cacggcgcga 2460
agcaaaaatt acggctcctc gctgcagacc tgcgagcagg gaaacgctcc cctcacagac 2520
gcgttgaatt gtccccacgc cgcgcccctg tagagaaata taaaaggtta ggatttgcca 2580
ctgaggttct tctttcatat acttccttt aaaatcttgc taggatacag ttctcacatc 2640
acatccgaac ataaacaacc atgcatgggt aaggaaaaga ctcacgtttc gaggccgcga 2700
ttaaattcca acatggatgc tgatttatat gggtataaat gggctcgcga taatgtcggg 2760
caatcaggtg cgacaatcta tcgattgtat gggaagcccg atgcgccaga gttgtttctg 2820
aaacatggca aaggtagcgt tgccaatgat gttacagatg agatggtcag actaaactgg 2880
ctgacggaat ttatgcctct tccgaccatc aagcatttta tccgtactcc tgatgatgca 2940
tggttactca ccactgcgat ccccggcaaa acagcattcc aggtattaga agaatatcct 3000
gattcaggtg aaaatattgt tgatgcgctg gcagtgttcc tgcgccggtt gcattcgatt 3060
cctgtttgta attgtccttt taacagcgat cgcgtatttc gcctcgctca ggcgcaatca 3120
cgaatgaata acggtttggt tgatgcgagt gattttgatg acgagcgtaa tggctggcct 3180
gttgaacaag tctggaaaga aatgcataag cttttgccat tctcaccgga ttcagtcgtc 3240
actcatggtg atttctcact tgataacctt atttttgacg aggggaaatt aataggttgt 3300
attgatgttg gacgagtcgg aatcgcagac cgataccagg atcttgccat cctatggaac 3360
tgcctcggtg agttttctcc ttcattacag aaacggcttt ttcaaaaata tggtattgat 3420
aatcctgata tgaataaatt gcagtttcat ttgatgctcg atgagttttt ctaatcagta 3480
ctgacaataa aaagattctt gttttcaaga acttgtcatt tgtatagttt ttttatattg 3540
tagttgttct attttaatca aatgttagcg tgatttatat ttttttcgc ctcgacatca 3600
tctgcccaga tgcgaagtta agtgcgcaga aagtaatatc atgcgtcaat cgtatgtgaa 3660
tgctggtcgc tatactgctg tcgattcgat actaacgccg ccatccagtg tcgaaaacga 3720
gctcgaattc atcgatgata tcagatccac tagtggccta tgcgaccgcg gatctgccgg 3780
tctccctata gtgagtcgta ttaatttcga taagccaggt taacctgcat taatgaatcg 3840
gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg 3900
actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agcatcgatg aattccacgg 3960
actatagact atactagtat actccgtcta ctgtacgata cacttccgct caggtccttg 4020
tcctttaacg aggccttacc actcttttgt tactctattg atccagctca gcaaaggcag 4080
tgtgatctaa gattctatct tcgcgatgta gtaaaactag ctagaccgag aaagagacta 4140
gaaatgcaaa aggcacttct acaatggctg ccatcattat tatccgatgt gacgctgcag 4200
cttctcaatg atattcgaat acgctttgag gagatacagc ctaatatccg acaaactgtt 4260
ttacagattt acgatcgtac ttgttaccca tcattgaatt ttgaacatcc gaacctggga 4320
gttttccctg aaacagatag tatatttgaa cctgtataat aatatatagt ctagcgcttt 4380
acggaagaca atgtatgtat ttcggttcct ggagaaacta ttgcatctat tgcataggta 4440
atcttgcacg tcgcatcccc ggttcatttt ctgcgtttcc atcttgcact tcaatagcat 4500
atctttgtta acgaagcatc tgtgcttcat tttgtagaac aaaaatgcaa cgcgagagcg 4560
ctaatttttc aaacaaagaa tctgagctgc atttttacag aacagaaatg caacgcgaaa 4620
gcgctatttt accaacgaag aatctgtgct tcatttttgt aaaacaaaaa tgcaacgcga 4680
gagcgctaat ttttcaaaca aagaatctga gctgcatttt tacagaacag aaatgcaacg 4740
cgagagcgct attttaccaa caaagaatct atacttcttt tttgttctac aaaaatgcat 4800
cccgagagcg ctatttttct aacaaagcat cttagattac tttttttctc ctttgtgcgc 4860

tctataatgc agtctcttga taacttttg cactgtaggt ccgttaaggt tagaagaagg 4920
ctactttggt gtctattttc tcttccataa aaaaagcctg actccacttc ccgcgtttac 4980
tgattactag cgaagctgcg ggtgcatttt ttcaagataa aggcatcccc gattatattc 5040
tataccgatg tggattgcgc atactttgtg aacagaaagt gatagcgttg atgattcttc 5100
attggtcaga aaattatgaa cggtttcttc tattttgtct ctatatacta cgtataggaa 5160
atgtttacat tttcgtattg ttttcgattc actctatgaa tagttcttac tacaattttt 5220
ttgtctaaag agtaatacta gagataaaca taaaaaatgt agaggtcgag tttagatgca 5280
agttcaagga gcgaaaggtg gatgggtagg ttatataggg atatagcaca gagatatata 5340
gcaaagagat actttgagc aatgtttgtg gaagcggtat tcgcaatatt ttagtagctc 5400
gttacagtcc ggtgcgtttt tggttttttg aaagtgcgtc ttcagagcgc ttttggtttt 5460
caaaagcgct ctgaagttcc tatactttct agagaatagg aacttcggaa taggaacttc 5520
aaagcgtttc cgaaaacgag cgcttccgaa aatgcaacgc gagctgcgca catacagctc 5580
actgttcacg tcgcacctat atctgcgtgt tgcctgtata tatatataca tgagaagaac 5640
ggcatagtgc gtgtttatgc ttaaatgcgt acttatatgc gtctatttat gtaggatgaa 5700
aggtagtcta gtacctcctg tgatattatc ccattccatg cggggtatcg tatgcttcct 5760
tcagcactac cctttagctg ttctatatgc tgccactcct caattggatt agtctcatcc 5820
ttcaatgcta tcatttcctt tgatattgga tcatatgcat agtaccgaga aactagtgcg 5880
aagtagtgat caggtattgc tgttatctga tgagtatacg ttgtcctggc cacggcagaa 5940
gcacgcttat cgctccaatt tcccacaaca ttagtcaact ccgttaggcc cttcattgaa 6000
agaaatgagg tcatcaaatg tcttccaatg tgagattttg ggccattttt tatagcaaag 6060
attgaataag gcgcattttt cttcaaagct ttattgtacg atctgactaa gttatctttt 6120
aataattggt attcctgttt attgcttgaa gaattgccgg tcctatttac tcgttttagg 6180
actggttcag aattcatcga tgctcactca aaggtcggta atacggttat ccacagaatc 6240
aggggataac gcaggaaaga ac 6262

<210> 243

<211> 6406
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS9923

<400> 243
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca 780
tttccccgaa aagtgccacc tgacgtccga tcaaaaatca tcgcttcgct gattaattac 840
cccagaaata aggctaaaaa actaatcgca ttatcatcct atggttgtta atttgattcg 900
ttcatttgaa ggtttgtggg gccaggttac tgccaatttt tcctcttcat aaccataaaa 960
gctagtattg tagaatcttt attgttcgga gcagtgcggc gcgaggcaca tctgcgtttc 1020
aggaacgcga ccggtgaaga cgaggacgca cggaggagag tcttccttcg gagggctgtc 1080
acccgctcgg cggcttctaa tccgtacttc aatatagcaa tgagcagtta agcgtattac 1140
tgaaagttcc aaagagaagg ttttttagg ctaatcgacc tcgagcagat ccgccaggcg 1200
tgtatatagc gtggatggcc aggcaacttt agtgctgaca catacaggca tatatatatg 1260
tgtgcgacga cacatgatca tatggcatgc atgtgctctg tatgtatata aaactcttgt 1320
tttcttcttt tctctaaata ttctttccttt atacattagg tcctttgtag cataaattac 1380
tatacttcta tagacacgca aacacaaata cacagcggcc ttgccaccat gggatccacc 1440
aagaccaaca gcgagattct ggaacagctg aaacaggcca gcgatggctt gctgttcatg 1500
tccgaaagcg agtacccgtt tgaagtcttc ttgtgggaag gttccgctcc accagtgacc 1560
catgagatcg tgttgcagca gactggtcat ggccaggatg ccccgttcaa agttgtggac 1620
atcgactcct tctttagcag agccaccact ccacaggatt ggtacgaaga tgaagagaac 1680
gccgtggttg ccaagtttca gaaactgctg gaagtcatca agtccaacct gaagaacccg 1740
caggtgtaca gattgggcga agtcgaactg gatgtgtacg tgattggcga gactccggct 1800
ggcaacctgg ccggcatcag caccaaagtt gtggagacct gacggccgac tgataactcg 1860
agcgatcctc tagacgagct cctcgagcct gcagcagctg aagctttgga cttcttcgcc 1920

agaggtttgg tcaagtctcc aatcaaggtt gtcggcttgt ctaccttgcc agaaatttac 1980
gaaaagatgg aaaagggtca aatcgttggt agatacgttg ttgacacttc taaataagcg 2040
aatttcttat gatttatgat ttttattatt aaataagtta taaaaaaaat aagtgtatac 2100
aaatttttaaa gtgactctta ggttttaaaa cgaaaattct tattcttgag taactctttc 2160
ctgtaggtca ggttgctttc tcaggtatag catgaggtcg ctcttattga ccacacctct 2220
accggcatgc aagcttggcg taatcatggt catagctgtt tcctgtgtga aattgttatc 2280
cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc tggggtgcct 2340
aatgagtgag ctaactcaca ttaattgcgt tgcgctcact gcccgctttc cagtcgggaa 2400
acctgtcgtg ccagcagatc tgtttagctt gcctcgtccc cgccgggtca cccggccagc 2460
gacatggagg cccagaatac cctccttgac agtcttgacg tgcgcagctc aggggcatga 2520
tgtgactgtc gcccgtacat ttagcccata catccccatg tataatcatt tgcatccata 2580
cattttgatg gccgcacggc gcgaagcaaa aattacggct cctcgctgca gacctgcgag 2640
cagggaaacg ctcccctcac agacgcgttg aattgtcccc acgccgcgcc cctgtagaga 2700
aatataaaag gttaggattt gccactgagg ttcttctttc atatacttcc ttttaaaatc 2760
ttgctaggat acagttctca catcacatcc gaacataaac aaccatgcat gggtaaggaa 2820
aagactcacg tttcgaggcc gcgattaaat tccaacatgg atgctgattt atatgggtat 2880
aaatgggctc gcgataatgt cgggcaatca ggtgcgacaa tctatcgatt gtatgggaag 2940
cccgatgcgc cagagttgtt tctgaaacat ggcaaaggta gcgttgccaa tgatgttaca 3000
gatgagatgg tcagactaaa ctggctgacg gaatttatgc ctcttccgac catcaagcat 3060
tttatccgta ctcctgatga tgcatggtta ctcaccactg cgatccccgg caaaacagca 3120
ttccaggtat tagaagaata tcctgattca ggtgaaaata ttgttgatgc gctggcagtg 3180
ttcctgcgcc ggttgcattc gattcctgtt tgtaattgtc cttttaacag cgatcgcgta 3240
tttcgcctcg ctcaggcgca atcacgaatg aataacggtt tggttgatgc gagtgatttt 3300
gatgacgagc gtaatggctg gcctgttgaa caagtctgga aagaaatgca taagcttttg 3360
ccattctcac cggattcagt cgtcactcat ggtgatttct cacttgataa ccttatttt 3420
gacgagggga aattaatagg ttgtattgat gttggacgag tcggaatcgc agaccgatac 3480
caggatcttg ccatcctatg gaactgcctc ggtgagtttt ctccttcatt acagaaacgg 3540
ctttttcaaa aatatggtat tgataatcct gatatgaata aattgcagtt tcatttgatg 3600
ctcgatgagt ttttctaatc agtactgaca ataaaaagat tcttgttttc aagaacttgt 3660
catttgtata gtttttttat attgtagttg ttctatttta atcaaatgtt agcgtgattt 3720
atatttttt tcgcctcgac atcatctgcc cagatgcgaa gttaagtgcg cagaaagtaa 3780
tatcatgcgt caatcgtatg tgaatgctgg tcgctatact gctgtcgatt cgatactaac 3840
gccgccatcc agtgtcgaaa acgagctcga attcatcgat gatatcagat ccactagtgg 3900
cctatgcgac cgcggatctg ccggtctccc tatagtgagt cgtattaatt tcgataagcc 3960
aggttaacct gcattaatga atcggccaac gcgcggggag aggcggtttg cgtattgggc 4020
gctcttccgc ttcctcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg 4080
tatcagcatc gatgaattcc acggactata gactatacta gtatactccg tctactgtac 4140
gatacacttc cgctcaggtc cttgtccttt aacgaggcct taccactctt ttgttactct 4200
attgatccag ctcagcaaag gcagtgtgat ctaagattct atcttcgcga tgtagtaaaa 4260
ctagctagac cgagaaagag actagaaatg caaaaggcac ttctacaatg gctgccatca 4320
ttattatccg atgtgacgct gcagcttctc aatgatattc gaatacgctt tgaggagata 4380
cagcctaata tccgacaaac tgttttacag atttacgatc gtacttgtta cccatcattg 4440
aattttgaac atccgaacct gggagttttc cctgaaacag atagtatatt tgaacctgta 4500
taataatata tagtctagcg ctttacggaa gacaatgtat gtatttcggt tcctggagaa 4560
actattgcat ctattgcata ggtaatcttg cacgtcgcat ccccggttca ttttctgcgt 4620
ttccatcttg cacttcaata gcatatcttt gttaacgaag catctgtgct tcattttgta 4680
gaacaaaaat gcaacgcgag agcgctaatt tttcaaacaa agaatctgag ctgcattttt 4740
acagaacaga aatgcaacgc gaaagcgcta ttttaccaac gaagaatctg tgcttcattt 4800
ttgtaaaaca aaaatgcaac gcgagagcgc taatttttca aacaaagaat ctgagctgca 4860
tttttacaga acagaaatgc aacgcgagag cgctatttta ccaacaaaga atctatactt 4920
ctttttgtt ctacaaaaat gcatcccgag agcgctattt ttctaacaaa gcatcttaga 4980
ttacttttt tctcctttgt gcgctctata atgcagtctc ttgataactt tttgcactgt 5040
aggtccgtta aggttagaag aaggctactt tggtgtctat tttctcttcc ataaaaaaag 5100
cctgactcca cttcccgcgt ttactgatta ctagcgaagc tgcgggtgca ttttttcaag 5160
ataaaggcat ccccgattat attctatacc gatgtggatt gcgcatactt tgtgaacaga 5220
aagtgatagc gttgatgatt cttcattggt cagaaaatta tgaacggttt cttctatttt 5280
gtctctatat actacgtata ggaaatgttt acattttcgt attgttttcg attcactcta 5340
tgaatagttc ttactacaat ttttttgtct aaagagtaat actagagata aacataaaaa 5400
atgtagaggt cgagtttaga tgcaagttca aggagcgaaa ggtggatggg taggttatat 5460
agggatatag cacagagata tatagcaaag agatactttt gagcaatgtt tgtggaagcg 5520
gtattcgcaa tattttagta gctcgttaca gtccggtgcg tttttggttt tttgaaagtg 5580
cgtcttcaga gcgcttttgg ttttcaaaag cgctctgaag ttcctatact ttctagagaa 5640
taggaacttc ggaataggaa cttcaaagcg tttccgaaaa cgagcgcttc cgaaaatgca 5700
acgcgagctg cgcacataca gctcactgtt cacgtcgcac ctatatctgc gtgttgcctg 5760
tatatatata tacatgagaa gaacggcata gtgcgtgttt atgcttaaat gcgtacttat 5820
atgcgtctat ttatgtagga tgaaaggtag tctagtacct cctgtgatat tatcccattc 5880
catgcggggt atcgtatgct tccttcagca ctacccttta gctgttctat atgctgccac 5940
tcctcaattg gattagtctc atccttcaat gctatcattt cctttgatat tggatcatat 6000

gcatagtacc gagaaactag tgcgaagtag tgatcaggta ttgctgttat ctgatgagta 6060
tacgttgtcc tggccacggc agaagcacgc ttatcgctcc aatttcccac aacattagtc 6120
aactccgtta ggcccttcat tgaaagaaat gaggtcatca aatgtcttcc aatgtgagat 6180
tttgggccat tttttatagc aaagattgaa taaggcgcat ttttcttcaa agctttattg 6240
tacgatctga ctaagttatc ttttaataat tggtattcct gtttattgct tgaagaattg 6300
ccggtcctat ttactcgttt taggactggt tcagaattca tcgatgctca ctcaaaggtc 6360
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaac 6406

<210> 244

<211> 11175
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS8590

<400> 244
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa atagggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg tggatctacg cacgctcggc 2460
tacagccagc agcaacagga gaagatcaaa ccgaaggttc gttcgacagt ggcgcagcac 2520
cacgaggcac tggtcggcca cgggtttaca cacgcgcaca tcgttgcgtt aagccaacac 2580
ccggcagcgt tagggaccgt cgctgtcaag tatcaggaca tgatcgcagc gttgccagag 2640
gcgacacacg aagcgatcgt tggcgtcggc aaacagtggt ccggcgcacg cgctctggag 2700
gccttgctca cggtggcggg agagttgaga ggtccaccgt tacagttgga cacaggccaa 2760
cttctcaaga ttgcaaaacg tggcggcgtg accgcagtgg aggcagtgca tgcatggcgc 2820
aatgcactga cgggtgcccc gctcaacttg accccggagc aggtggtggc catcgccagc 2880
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2940

```
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag   3000
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg   3060
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag   3120
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc   3180
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg   3240
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc   3300
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg   3360
accccggagc aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg   3420
gtgcaggcgc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg   3480
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg   3540
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat   3600
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac   3660
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg   3720
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag   3780
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg   3840
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc   3900
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag   3960
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag   4020
gcgctggaga cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg   4080
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag   4140
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc   4200
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg   4260
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc   4320
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg   4380
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg   4440
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg   4500
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg   4560
gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg tggccatcgc cagcaatggc   4620
ggcggcaggc cggcgctgga gagcattgtt gcccagttat ctcgccctga tccgggatcc   4680
tccggaccga acagaggcgt taccaagcaa ctagttaaat ccgaactgga agagaaaaag   4740
tccgagttgc gtcataaact gaagtacgtt ccacatgaat acattgaact gatcgagatt   4800
gcccgtaaca gcacccagga cagaatcctg gaaatgaaag tgatggagtt tttcatgaaa   4860
gtgtacggct acagaggcaa acatttgggc ggtagccgta aaccggatgg tgccatctac   4920
accgttggca gcccgatcga ttacggcgtg atcgttgata ccaaagccta ctccggtggc   4980
tacaacctgc cgattggcca ggctgacgaa atgcagagat acgtggaaga gaaccagact   5040
cgtaacaagc atatcaaccc gaacgaatgg tggaaagtct acccgagttc cgtgactgag   5100
ttcaaatttc tgtttgtgag cggccatttc aaaggcaact acaaagccca gctgaccaga   5160
ctgaaccata tcaccaactg caacggtgcc gtcttgagcg tggaagaact gttgattggt   5220
ggcgaaatga tcaaggctgg caccctgact ctggaggaag ttagacgtaa attcaacaac   5280
ggtgaaatca actttggatc ctcggccgac tgataactcg agcgatcctc tagacgagct   5340
cctcgagcct gcagcagctg aagctttgga cttcttcgcc agaggtttgg tcaagtctcc   5400
aatcaaggtt gtcggcttgt ctaccttgcc agaaatttac gaaaagatgg aaaagggtca   5460
aatcgttggt agatacgttg ttgacacttc taaataagcg aatttcttat gatttatgat   5520
ttttattatt aaataagtta taaaaaaaat aagtgtatac aaattttaaa gtgactctta   5580
ggttttaaaa cgaaaattct tattcttgag taactctttc ctgtaggtca ggttgctttc   5640
tcaggtatag catgaggtcg ctcttattga ccacacctct accggcatgc aagcttggcg   5700
taatcatggt catagctgtt tcctgtgtga aattgttatc cgctcacaat tccacacaac   5760
atacgagccg gaagcataaa gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca   5820
ttaattgcgt tgcgctcact gcccgctttc cagtcgggaa acctgtcgtg ccagcagatc   5880
tattacatta tgggtggtat gttggaataa aaatcaacta tcatctacta actagtattt   5940
acgttactag tatattatca tatacggtgt tagaagatga cgcaaatgat gagaaatagt   6000
catctaaatt agtggaagct gaaacgcaag gattgataat gtaataggat caatgaatat   6060
taacatataa aatgatgata ataatattta tagaattgtg tagaattgca gattcccttt   6120
tatggattcc taaatcctcg aggagaactt ctagtatatc tacataccta atattattgc   6180
cttattaaaa atggaatccc aacaattaca tcaaaatcca cattctcttc aaaatcaatt   6240
gtcctgtact tccttgttca tgtgtgttca aaaacgttat atttatagga taattatact   6300
ctatttctca acaagtaatt ggttgtttgg ccgagcggtc taaggcgcct gattcaagaa   6360
atatcttgac cgcagttaac tgtgggaata ctcaggtatc gtaagatgca agagttcgaa   6420
tctcttagca accattattt ttttcctcaa cataacgaga acacacaggg gcgctatcgc   6480
acagaatcaa attcgatgac tggaaatttt ttgttaattt cagaggtcgc ctgacgcata   6540
taccttttc aactgaaaaa ttgggagaaa aaggaaaggt gagagccgcg gaaccggctt   6600
ttcatataga atagagaagc gttcatgact aaatgcttgc atcacaatac ttgaagttga   6660
caatattatt taaggaccta ttgtttttc caataggtgg ttagcaatcg tcttactttc   6720
taacttttct tacctttac atttcagcaa tatatatata tatatttcaa ggatatacca   6780
ttctaatgtc tgcccctaag aagatcgtcg ttttgccagg tgaccacgtt ggtcaagaaa   6840
tcacagccga agccattaag gttcttaaag ctatttctga tgttcgttcc aatgtcaagt   6900
tcgatttcga aaatcattta attggtggtg ctgctatcga tgctacaggt gtcccacttc   6960
cagatgaggc gctggaagcc tccaagaagg ttgatgccgt tttgttaggt gctgtgggtg   7020
```

```
gtcctaaatg gggtaccggt agtgttagac ctgaacaagg tttactaaaa atccgtaaag   7080
aacttcaatt gtacgccaac ttaagaccat gtaactttgc atccgactct cttttagact   7140
tatctccaat caagccacaa tttgctaaag gtactgactt cgttgttgtc agagaattag   7200
tgggaggtat ttactttggt aagagaaagg aagacgatgg tgatggtgtc gcttgggata   7260
gtgaacaata caccgttcca gaagtgcaaa gaatcacaag aatggccgct ttcatggccc   7320
tacaacatga gccaccattg cctatttggt ccttggataa agctaatgtt ttggcctctt   7380
caagattatg gagaaaaact gtggaggaaa ccatcaagaa cgaattccct acattgaagg   7440
ttcaacatca attgattgat tctgccgcca tgatcctagt taagaaccca acccacctaa   7500
atggtattat aatcaccagc aacatgtttg gtgatatcat ctccgatgaa gcctccgtta   7560
tcccaggttc cttgggtttg ttgccatctg cgtccttggc ctctttgcca gacaagaaca   7620
ccgcatttgg tttgtacgaa ccatgccacg gttctgctcc agatttgcca aagaataagg   7680
tcaaccctat cgccactatc ttgtctgctg caatgatgtt gaaattgtca ttgaacttgc   7740
ctgaagaagg taaggccatt gaagatgcag ttaaaaaggt tttggatgca ggtatcagaa   7800
ctggtgattt aggtggttcc aacagtacca cggaagtcgg tgatgctgtc gccgaagaag   7860
ttaagaaaat ccttgcttaa aaagattctc ttttttatg atatttgtac ataaacttta   7920
taaatgaaat tcataatага aacgacacga aattacaaaa tggaatatgt tcatagggta   7980
gacgaaacta tatacgcaat ctacatacat ttatcaagaa ggagaaaaag gaggatgtaa   8040
aggaatacag gtaagcaaat tgatactaat ggctcaacgt gataaggaaa aagaattgca   8100
ctttaacatt aatattgaca aggaggaggg caccacacaa aaagttaggt gtaacagaaa   8160
atcatgaaac tatgattcct aatttatata ttggaggatt ttctctaaaa aaaaaaaaat   8220
acaacaaata aaaaacactc aatgacctga ccatttgatg gagtttaagt caatаccttc   8280
ttgaaccatt tcccataatg gtgaaagttc cctcaagaat tttactctgt cagaaacggc   8340
cttaacgacg tagtcgacct cctcttcagt actaaatcta ccaataccaa atctgatgga   8400
agaatgggct aatgcatcat ccttacccag cgcatgtaaa acataagaag gttctaggga   8460
agcagatgta caggctgaac ccgaggataa tgcgatatcc cttagtgcca tcaataaaga   8520
ttctccttcc acgtaggcga aagaaacgtt aacacaccct ggataacgat gatctggaga   8580
tccgttcaac gtggtatgtt cagcggataa tagacctttg actaatttat cggatagtct   8640
tttgatgtga gcttggtcgt tgtcaaattc tttcttcatc aatctcgcag cttcaccaaa   8700
tcccgctacc aatgggggggg ccaaagtacc agatctgctg cattaatgaa tcggccaacg   8760
cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct   8820
gcgctcggtc gttcggctgc ggcgagcggt atcagcatcg atgaattcca cggactatag   8880
actatactag tatactccgt ctactgtacg atacacttcc gctcaggtcc ttgtccttta   8940
acgaggcctt accactcttt tgttactcta ttgatccagc tcagcaaagg cagtgtgatc   9000
taagattcta tcttcgcgat gtagtaaaac tagctagacc gagaaagaga ctagaaatgc   9060
aaaaggcact tctacaatgg ctgccatcat tattatccga tgtgacgctg cagcttctca   9120
atgatattcg aatacgcttt gaggagatac agcctaatat ccgacaaact gttttacaga   9180
tttacgatcg tacttgttac ccatcattga attttgaaca tccgaacctg ggagttttcc   9240
ctgaaacaga tagtatattt gaacctgtat aataatatat agtctagcgc tttacggaag   9300
acaatgtatg tatttcggtt cctggagaaa ctattgcatc tattgcatag gtaatcttgc   9360
acgtcgcatc cccggttcat tttctgcgtt tccatcttgc acttcaatag catatctttg   9420
ttaacgaagc atctgtgctt cattttgtag aacaaaaatg caacgcgaga gcgctaattt   9480
ttcaaacaaa gaatctgagc tgcattttta cagaacagaa atgcaacgcg aaagcgctat   9540
tttaccaacg aagaatctgt gcttcatttt tgtaaaacaa aaatgcaacg cgagagcgct   9600
aatttttcaa acaaagaatc tgagctgcat ttttacagaa cagaaatgca acgcgagagc   9660
gctattttac caacaaagaa tctatacttc ttttttgttc tacaaaaatg catcccgaga   9720
gcgctatttt tctaacaaag catcttagat tacttttttt ctcctttgtg cgctctataa   9780
tgcagtctct tgataacttt ttgcactgta ggtccgttaa ggttagaaga aggctacttt   9840
ggtgtctatt ttctcttcca taaaaaaagc ctgactccac ttcccgcgtt tactgattac   9900
tagcgaagct gcgggtgcat tttttcaaga taaaggcatc cccgattata ttctataccg   9960
atgtggattg cgcatacttt gtgaacagaa agtgatagcg ttgatgattc ttcattggtc  10020
agaaaattat gaacggtttc ttctattttg tctctatata ctacgtatag gaaatgttta  10080
cattttcgta ttgttttcga ttcactctat gaatagttct tactacaatt tttttgtcta  10140
aagagtaata ctagagataa acataaaaaa tgtagaggtc gagtttagat gcaagttcaa  10200
ggagcgaaag gtggatgggt aggttatata gggatatagc acagagatat atagcaaaga  10260
gatacttttg agcaatgttt gtggaagcgg tattcgcaat attttagtag ctcgttacag  10320
tccggtgcgt ttttggtttt ttgaaagtgc gtcttcagag cgcttttggt tttcaaaagc  10380
gctctgaagt tcctatactt tctagagaat aggaacttcg gaataggaac ttcaaagcgt  10440
ttccgaaaac gagcgcttcc gaaaatgcaa cgcgagctgc gcacatacag ctcactgttc  10500
acgtcgcacc tatatctgcg tgttgcctgt atatatatat acatgagaag aacggcatag  10560
tgcgtgttta tgcttaaatg cgtacttata tgcgtctatt tatgtaggat gaaaggtagt  10620
ctagtacctc ctgtgatatt atcccattcc atgcggggta tcgtatgctt ccttcagcac  10680
taccctttag ctgttctata tgctgccact cctcaattgg attagtctca tccttcaatg  10740
ctatcatttc ctttgatatt ggatcatatg catagtaccg agaaactagt gcgaagtagt  10800
gatcaggtat tgctgttatc tgatgagtat acgttgtcct ggccacggca gaagcacgct  10860
tatcgctcca atttcccaca acattagtca actccgttag gcccttcatt gaaagaaatg  10920
aggtcatcaa atgtcttcca atgtgagatt ttgggccatt ttttatagca aagattgaat  10980
aaggcgcatt tttcttcaaa gctttattgt acgatctgac taagttatct tttaataatt  11040
ggtattcctg tttattgctt gaagaattgc cggtcctatt tactcgtttt aggactggtt  11100
```

cagaattcat cgatgctcac tcaaaggtcg gtaatacggt tatccacaga atcaggggat 11160
aacgcaggaa agaac 11175

<210> 245

<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> RagT2-R target

<400> 245
tgtttatggt tacttat 17

<210> 246

<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> NptIIT5-L target

<400> 246
tccttgcgca gctgtgc 17

<210> 247

<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> NptIIT5-R target

<400> 247
tagcagccag tcccttc 17

<210> 248

<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> NptIIT6-L target

<400> 248
tgaagcggga agggact 17

<210> 249

<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> NptIIT6-R target

<400> 249
tgacaggaga tcctgcc 17

<210> 250

<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> RagT2.1 target

<400> 250
tatatttaag cacttatatg tgtgtaacag gtataagtaa ccataaaca 49

<210> 251

<211> 55
<212> DNA
<213> Artificial sequence

<220>
<223> NptIIT5.1 target

<400> 251
tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc tgcta 55

<210> 252

<211> 58
<212> DNA
<213> Artificial sequence

<220>
<223> NptIIT6.1 target

<400> 252
tgaagcggga agggactggc tgctattggg cgaagtgccg gggcaggatc tcctgtca 58

<210> 253

<211> 530
<212> PRT
<213> Artificial sequence

<220>
<223> RagT2-R RVD insert

<400> 253
Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys
1 5 10 15

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
20 25 30

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
35 40 45

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
50 55 60

Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
65 70 75 80

```
Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
                85                  90                  95

Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
            100                 105                 110

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
        115                 120                 125

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
    130                 135                 140

Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala
145                 150                 155                 160

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
                165                 170                 175

Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val
            180                 185                 190

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln
        195                 200                 205

Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu
    210                 215                 220

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
225                 230                 235                 240

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala
                245                 250                 255

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
            260                 265                 270

Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys
        275                 280                 285

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
    290                 295                 300

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
305                 310                 315                 320

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
                325                 330                 335

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn
            340                 345                 350
```

```
Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val
        355                 360                 365

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
    370                 375                 380

Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
385                 390                 395                 400

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
                405                 410                 415

Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
            420                 425                 430

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
        435                 440                 445

Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val
    450                 455                 460

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
465                 470                 475                 480

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu
                485                 490                 495

Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
            500                 505                 510

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala
        515                 520                 525

Leu Glu
    530
```

<210> 254

<211> 530
<212> PRT
<213> Artificial sequence

<220>
<223> NptIIT5-L RVD insert

<400> 254

```
Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
1               5                   10                  15

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
            20                  25                  30
```

```
His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly
        35                  40                  45

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
    50                  55                  60

Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
65                  70                  75                  80

Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
                85                  90                  95

Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
            100                 105                 110

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
        115                 120                 125

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
    130                 135                 140

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
145                 150                 155                 160

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
                165                 170                 175

Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val
            180                 185                 190

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln
        195                 200                 205

Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu
    210                 215                 220

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
225                 230                 235                 240

Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala
                245                 250                 255

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
            260                 265                 270

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys
        275                 280                 285

Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala
    290                 295                 300
```

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly
305 310 315 320

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
325 330 335

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His
340 345 350

Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
355 360 365

Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
370 375 380

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
385 390 395 400

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
405 410 415

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
420 425 430

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
435 440 445

Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val
450 455 460

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln
465 470 475 480

Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu
485 490 495

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
500 505 510

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala
515 520 525

Leu Glu
530

<210> 255

<211> 530
<212> PRT
<213> Artificial sequence

<220>
<223> NptIIT5-R RVD insert

<400> 255
Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys
1 5 10 15

Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala
20 25 30

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly
35 40 45

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
50 55 60

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His
65 70 75 80

Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
85 90 95

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
100 105 110

Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu
115 120 125

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
130 135 140

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
145 150 155 160

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
165 170 175

Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val
180 185 190

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
195 200 205

Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu
210 215 220

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
225 230 235 240

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala
245 250 255

Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly
260 265 270

```
Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys
        275                 280                 285

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
    290                 295                 300

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
305                 310                 315                 320

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
                325                 330                 335

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His
            340                 345                 350

Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
        355                 360                 365

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
    370                 375                 380

Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
385                 390                 395                 400

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
                405                 410                 415

Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
            420                 425                 430

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
        435                 440                 445

Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val
    450                 455                 460

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln
465                 470                 475                 480

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu
                485                 490                 495

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
            500                 505                 510

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala
        515                 520                 525

Leu Glu
    530
```

<210> 256

<211> 530
<212> PRT
<213> Artificial sequence

<220>
<223> NptIIT6-L RVD insert

<400> 256
Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys
1   5   10   15

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
            20   25   30

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly
        35   40   45

Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys
    50   55   60

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn
65   70   75   80

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val
                85   90   95

Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
            100   105   110

Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
        115   120   125

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
    130   135   140

Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
145   150   155   160

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
                165   170   175

Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val
            180   185   190

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln
        195   200   205

Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu
    210   215   220

```
Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
225                 230                 235                 240

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala
                245                 250                 255

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
            260                 265                 270

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys
        275                 280                 285

Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala
    290                 295                 300

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly
305                 310                 315                 320

Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys
                325                 330                 335

Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
            340                 345                 350

Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
        355                 360                 365

Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
    370                 375                 380

Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
385                 390                 395                 400

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
                405                 410                 415

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
            420                 425                 430

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
        435                 440                 445

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val
    450                 455                 460

Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
465                 470                 475                 480

Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu
                485                 490                 495
```

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
500 505 510

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala
515 520 525

Leu Glu
530

<210> 257

<211> 530
<212> PRT
<213> Artificial sequence

<220>
<223> NptIIT6-R RVD insert

<400> 257
Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys
1 5 10 15

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
20 25 30

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly
35 40 45

Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys
50 55 60

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His
65 70 75 80

Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
85 90 95

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
100 105 110

Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu
115 120 125

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
130 135 140

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
145 150 155 160

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
165 170 175

Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val
180 185 190

```
Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
        195                 200                 205

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu
    210                 215                 220

Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
225                 230                 235                 240

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala
                245                 250                 255

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
            260                 265                 270

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys
        275                 280                 285

Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala
    290                 295                 300

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
305                 310                 315                 320

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
                325                 330                 335

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His
            340                 345                 350

Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
        355                 360                 365

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
    370                 375                 380

Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
385                 390                 395                 400

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
                405                 410                 415

Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
            420                 425                 430

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
        435                 440                 445

Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val
    450                 455                 460
```

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
465 470 475 480

Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu
485 490 495

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
500 505 510

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala
515 520 525

Leu Glu
530

<210> 258

<211> 9183
<212> DNA
<213> Artificial sequence

<220>
<223> TevCreD05 in pCLS6618

<400> 258
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa atagggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2160

agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg ccaagtctgg catctaccag 2460
attaagaaca ccctgaataa caaagtgtat gttggatctg ccaaggactt tgagaaaaga 2520
tggaagaggc acttcaagga tctggaaaag ggttgccaca gctccatcaa gctccaaagg 2580
tctttcaaca aacatggcaa tgtctttgag tgctccattt tggaagagat cccttatgaa 2640
aaggacctca ttatagagag ggaaaacttc tggatcaagg agctgaactc aaaaatcaat 2700
ggatacaaca ttgcagatgc cacatttggt gacacctgtt ctactcaccc ccttaaggaa 2760
gagatcatta aaaagaggtc tgagactgtg aaggctaaga tgttgaaaca aggtccatct 2820
ggcaatacca aatataacaa agagttcctg ctgtacctgg ccggctttgt ggacggtaac 2880
ggtagcatca tcgctcagat taaaccaaac cagtcttata agtttaaaca tcagctaagc 2940
ttgacctttc aggtgactca aaagacccag cgccgttggt ttctggacaa actagtggat 3000
gaaattggcg ttggttacgt acgtgatcgc ggatccgttt ccgattacat tttaagcgaa 3060
atcaagccgc tgcacaactt cctgactcaa ctgcagccgt ttctgaaact gaaacagaaa 3120
caggcaaacc tggttctgaa aattatcgaa cagctgccgt ctgcaaaaga atccccggac 3180
aaattcctgg aagtttgtac ctgggtggat cagattgcag ctctgaacga ttctaagacg 3240
cgtaaaacca cttctgaaac cgttcgtgct gtgctggaca gcctgagcga gaagaagaaa 3300
tcctccccgg cggccgactg ataactcgag cgatcctcta gacgagctcc tcgagcctgc 3360
agcagctgaa gctttggact tcttcgccag aggtttggtc aagtctccaa tcaaggttgt 3420
cggcttgtct accttgccag aaatttacga aaagatggaa aagggtcaaa tcgttggtag 3480
atacgttgtt gacacttcta aataagcgaa tttcttatga tttatgattt ttattattaa 3540
ataagttata aaaaaaataa gtgtatacaa attttaaagt gactcttagg ttttaaaacg 3600
aaaattctta ttcttgagta actctttcct gtaggtcagg ttgctttctc aggtatagca 3660
tgaggtcgct cttattgacc acacctctac cggcatgcaa gcttggcgta atcatggtca 3720
tagctgtttc ctgtgtgaaa ttgttatccg ctcacaattc cacacaacat acgagccgga 3780
agcataaagt gtaaagcctg gggtgcctaa tgagtgagct aactcacatt aattgcgttg 3840
cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agcagatcta ttacattatg 3900
ggtggtatgt tggaataaaa atcaactatc atctactaac tagtatttac gttactagta 3960
tattatcata tacggtgtta gaagatgacg caaatgatga gaaatagtca tctaaattag 4020
tggaagctga aacgcaagga ttgataatgt aataggatca atgaatatta acatataaaa 4080
tgatgataat aatatttata gaattgtgta gaattgcaga ttccctttta tggattccta 4140
aatcctcgag gagaacttct agtatatcta catacctaat attattgcct tattaaaaat 4200
ggaatcccaa caattacatc aaaatccaca ttctcttcaa aatcaattgt cctgtacttc 4260
cttgttcatg tgtgttcaaa aacgttatat ttataggata attatactct atttctcaac 4320
aagtaattgg ttgtttggcc gagcggtcta aggcgcctga ttcaagaaat atcttgaccg 4380
cagttaactg tgggaatact caggtatcgt aagatgcaag agttcgaatc tcttagcaac 4440
cattattttt ttcctcaaca taacgagaac acacaggggc gctatcgcac agaatcaaat 4500
tcgatgactg gaaatttttt gttaatttca gaggtcgcct gacgcatata ccttttttcaa 4560
ctgaaaaatt gggagaaaaa ggaaaggtga gagccgcgga accggctttt catatagaat 4620
agagaagcgt tcatgactaa atgcttgcat cacaatactt gaagttgaca atattattta 4680
aggacctatt gtttttttca ataggtggtt agcaatcgtc ttactttcta acttttctta 4740
ccttttacat ttcagcaata tatatatata tatttcaagg atataccatt ctaatgtctg 4800
cccctaagaa gatcgtcgtt ttgccaggtg accacgttgg tcaagaaatc acagccgaag 4860
ccattaaggt tcttaaagct atttctgatg ttcgttccaa tgtcaagttc gatttcgaaa 4920
atcatttaat tggtggtgct gctatcgatg ctacaggtgt cccacttcca gatgaggcgc 4980
tggaagcctc caagaaggtt gatgccgttt tgttaggtgc tgtgggtggt cctaaatggg 5040
gtaccggtag tgttagacct gaacaaggtt tactaaaaat ccgtaaagaa cttcaattgt 5100
acgccaactt aagaccatgt aactttgcat ccgactctct tttagactta tctccaatca 5160
agccacaatt tgctaaaggt actgacttcg ttgttgtcag agaattagtg ggaggtattt 5220
actttggtaa gagaaaggaa gacgatggtg atggtgtcgc ttgggatagt gaacaataca 5280
ccgttccaga agtgcaaaga atcacaagaa tggccgcttt catggcccta caacatgagc 5340
caccattgcc tatttggtcc ttggataaag ctaatgtttt ggcctcttca agattatgga 5400
gaaaaactgt ggaggaaacc atcaagaacg aattccctac attgaaggtt caacatcaat 5460
tgattgattc tgccgccatg atcctagtta agaacccaac ccacctaaat ggtattataa 5520
tcaccagcaa catgtttggt gatatcatct ccgatgaagc ctccgttatc ccaggttcct 5580
tgggtttgtt gccatctgcg tccttggcct ctttgccaga caagaacacc gcatttggtt 5640
tgtacgaacc atgccacggt tctgctccag atttgccaaa gaataaggtc aaccctatcg 5700
ccactatctt gtctgctgca atgatgttga aattgtcatt gaacttgcct gaagaaggta 5760
aggccattga agatgcagtt aaaaaggttt tggatgcagg tatcagaact ggtgatttag 5820
gtggttccaa cagtaccacg gaagtcggtg atgctgtcgc cgaagaagtt aagaaaatcc 5880
ttgcttaaaa agattctctt tttttatgat atttgtacat aaactttata aatgaaattc 5940
ataatagaaa cgacacgaaa ttacaaaatg gaatatgttc atagggtaga cgaaactata 6000
tacgcaatct acatacattt atcaagaagg agaaaaagga ggatgtaaag gaatacaggt 6060
aagcaaattg atactaatgg ctcaacgtga taaggaaaaa gaattgcact ttaacattaa 6120
tattgacaag gaggagggca ccacacaaaa agttaggtgt aacagaaaat catgaaacta 6180
tgattcctaa tttatatatt ggaggatttt ctctaaaaaa aaaaaaatac aacaaataaa 6240

aaacactcaa tgacctgacc atttgatgga gtttaagtca ataccttctt gaaccatttc 6300
ccataatggt gaaagttccc tcaagaattt tactctgtca gaaacggcct taacgacgta 6360
gtcgacctcc tcttcagtac taaatctacc aataccaaat ctgatggaag aatgggctaa 6420
tgcatcatcc ttacccagcg catgtaaaac ataagaaggt tctagggaag cagatgtaca 6480
ggctgaaccc gaggataatg cgatatccct tagtgccatc aataaagatt ctccttccac 6540
gtaggcgaaa gaaacgttaa cacaccctgg ataacgatga tctggagatc cgttcaacgt 6600
ggtatgttca gcggataata gacctttgac taatttatcg gatagtcttt tgatgtgagc 6660
ttggtcgttg tcaaattctt tcttcatcaa tctcgcagct tcaccaaatc ccgctaccaa 6720
tgggggggcc aaagtaccag atctgctgca ttaatgaatc ggccaacgcg cggggagagg 6780
cggtttgcgt attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt 6840
tcggctgcgg cgagcggtat cagcatcgat gaattccacg gactatagac tatactagta 6900
tactccgtct actgtacgat acacttccgc tcaggtcctt gtcctttaac gaggccttac 6960
cactcttttg ttactctatt gatccagctc agcaaaggca gtgtgatcta agattctatc 7020
ttcgcgatgt agtaaaacta gctagaccga gaaagagact agaaatgcaa aaggcacttc 7080
tacaatggct gccatcatta ttatccgatg tgacgctgca gcttctcaat gatattcgaa 7140
tacgctttga ggagatacag cctaatatcc gacaaactgt tttacagatt tacgatcgta 7200
cttgttaccc atcattgaat tttgaacatc cgaacctggg agttttccct gaaacagata 7260
gtatatttga acctgtataa taatatatag tctagcgctt tacggaagac aatgtatgta 7320
tttcggttcc tggagaaact attgcatcta ttgcataggt aatcttgcac gtcgcatccc 7380
cggttcattt tctgcgtttc catcttgcac ttcaatagca tatctttgtt aacgaagcat 7440
ctgtgcttca ttttgtagaa caaaaatgca acgcgagagc gctaattttt caaacaaaga 7500
atctgagctg catttttaca gaacagaaat gcaacgcgaa agcgctattt taccaacgaa 7560
gaatctgtgc ttcatttttg taaaacaaaa atgcaacgcg agagcgctaa tttttcaaac 7620
aaagaatctg agctgcattt ttacagaaca gaaatgcaac gcgagagcgc tattttacca 7680
acaaagaatc tatacttctt ttttgttcta caaaaatgca tcccgagagc gctatttttc 7740
taacaaagca tcttagatta ctttttttct cctttgtgcg ctctataatg cagtctcttg 7800
ataacttttt gcactgtagg tccgttaagg ttagaagaag gctactttgg tgtctatttt 7860
ctcttccata aaaaaagcct gactccactt cccgcgttta ctgattacta gcgaagctgc 7920
gggtgcattt tttcaagata aaggcatccc cgattatatt ctataccgat gtggattgcg 7980
catactttgt gaacagaaag tgatagcgtt gatgattctt cattggtcag aaaattatga 8040
acggtttctt ctattttgtc tctatatact acgtatagga aatgtttaca ttttcgtatt 8100
gttttcgatt cactctatga atagttctta ctacaatttt tttgtctaaa gagtaatact 8160
agagataaac ataaaaaatg tagaggtcga gtttagatgc aagttcaagg agcgaaaggt 8220
ggatgggtag gttatatagg gatatagcac agagatatat agcaaagaga tacttttgag 8280
caatgtttgt ggaagcggta ttcgcaatat tttagtagct cgttacagtc cggtgcgttt 8340
ttggtttttt gaaagtgcgt cttcagagcg cttttggttt tcaaaagcgc tctgaagttc 8400
ctatactttc tagagaatag gaacttcgga ataggaactt caaagcgttt ccgaaaacga 8460
gcgcttccga aaatgcaacg cgagctgcgc acatacagct cactgttcac gtcgcaccta 8520
tatctgcgtg ttgcctgtat atatatatac atgagaagaa cggcatagtg cgtgtttatg 8580
cttaaatgcg tacttatatg cgtctattta tgtaggatga aaggtagtct agtacctcct 8640
gtgatattat cccattccat gcggggtatc gtatgcttcc ttcagcacta ccctttagct 8700
gttctatatg ctgccactcc tcaattggat tagtctcatc cttcaatgct atcatttcct 8760
ttgatattgg atcatatgca tagtaccgag aaactagtgc gaagtagtga tcaggtattg 8820
ctgttatctg atgagtatac gttgtcctgg ccacggcaga agcacgctta tcgctccaat 8880
ttcccacaac attagtcaac tccgttaggc ccttcattga aagaaatgag gtcatcaaat 8940
gtcttccaat gtgagatttt gggccatttt ttatagcaaa gattgaataa ggcgcatttt 9000
tcttcaaagc tttattgtac gatctgacta agttatcttt taataattgg tattcctgtt 9060
tattgcttga agaattgccg gtcctattta ctcgttttag gactggttca gaattcatcg 9120
atgctcactc aaaggtcggt aatacggtta tccacagaat caggggataa cgcaggaaag 9180
aac 9183

<210> 259

<211> 36
<212> DNA
<213> Artificial sequence

<220>
<223> CMP_G115

<400> 259
tggtatcggc cgatggacct tgtttcaaca tcttag 36

<210> 260

<211> 3866

<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15775

<400> 260
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggtg gatctacgca cgctcggcta cagccagcag 480
caacaggaga agatcaaacc gaaggttcgt tcgacagtgg cgcagcacca cgaggcactg 540
gtcggccacg ggtttacaca cgcgcacatc gttgcgttaa gccaacaccc ggcagcgtta 600
gggaccgtcg ctgtcaagta tcaggacatg atcgcagcgt tgccagaggc gacacacgaa 660
gcgatcgttg gcgtcggcaa acagtggtcc ggcgcacgcg ctctggaggc cttgctcacg 720
gtggcgggag agttgagagg tccaccgtta cagttggaca caggccaact tctcaagatt 780
gcaaaacgtg gcggcgtgac cgcagtggag gcagtgcatg catggcgcaa tgcactgacg 840
ggtgccccgc tcaacttgac cggagacgcc cgggggatca ggtcacgtgc gtctcggagc 900
attgttgccc agttatctcg ccctgatccg ggatccaagt ctggcatcta ccagattaag 960
aacaccctga ataacaaagt gtatgttgga tctgccaagg actttgagaa aagatggaag 1020
aggcacttca aggatctgga aaagggttgc cacagctcca tcaagctcca aaggtctttc 1080
aacaaacatg gcaatgtctt tgagtgctcc attttggaag agatccctta tgaaaaggac 1140
ctcattatag agagggaaaa cttctggatc aaggagctga actcaaaaat caatggatac 1200
aacattgcag atgccacatt tggtgacacc tgttctactc acccccttaa ggaagagatc 1260
attaaaaaga ggtctgagac tgtgaaggct aagatgttga aacaaggtcc atcggccgac 1320
tgataagtcg accaccacca ccaccaccac tgataagagc tcctcgagcc tgcagcagct 1380
gaagctttgg acttctggcg cgcccaattc gccctatagt gagtcgtatt acgtcgcgct 1440
cactggccgt cgttttacaa cgtcgtgact gggaaaaccc tggcgttacc caacttaatc 1500
gccttgcagc acatcccccc ttcgccagct ggcgtaatag cgaagaggcc cgcaccgaaa 1560
cgcccttccc aacagttgcg cagcctgaat ggcgaatggg agcgccctgt agcggcgcat 1620
taagcgcggc gggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc agcgccctag 1680
cgcccgctcc tttcgctttc ttcccttcct ttctcgccac gttcgccggc tttccccgtc 1740
aagctctaaa tcgggggctc cctttagggt tccgatttag tgctttacgg cacctcgacc 1800
ccaaaaaact tgattagggt gatggttggc ctgtagtggg ccatagccct gatagacggt 1860
ttttcgccct ttgacgttgg agtccacgtt ctttaatagt ggactcttgt tccaaactgg 1920
aacaacactc aaccctatct cggtctattc ttttgattta taagggattt tgccgatttc 1980
ggcctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt ttaacaaaat 2040
attaacgctt acaatttagg tggcactttt cggggaaatg tgcgcggaac ccctatttgt 2100
ttatttttct aaatacattc aaatatgtat ccgctcatga gacaataacc ctgataaatg 2160
cttcaataat attgaaaaag gaagagtatg agtattcaac atttccgtgt cgcccttatt 2220
ccctttttg cggcattttg ccttcctgtt tttgctcacc cagaaacgct ggtgaaagta 2280
aaagatgctg aagatcagtt gggtgcacga gtgggttaca tcgaactgga tctcaacagc 2340
ggtaagatcc ttgagagttt tcgccccgaa gaacgttttc caatgatgag cacttttaaa 2400
gttctgctat gtggcgcggt attatcccgt attgacgccg ggcaagagca actcggtcgc 2460
cgcatacact attctcagaa tgacttggtt gagtactcac cagtcacaga aaagcatctt 2520
acggatggca tgacagtaag agaattatgc agtgctgcca taaccatgag tgataacact 2580
gcggccaact tacttctgac aacgatcgga ggaccgaagg agctaaccgc ttttttgcac 2640
aacatggggg atcatgtaac tcgccttgat cgttgggaac cggagctgaa tgaagccata 2700
ccaaacgacg agcgtgacac cacgatgcct gtagcaatgg caacaacgtt gcgcaaacta 2760
ttaactggcg aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg 2820
gataaagttg caggaccact tctgcgctcg gcccttccgg ctggctggtt tattgctgat 2880
aaatctggag ccggtgagcg tggttctcgc ggtatcattg cagcactggg gccagatggt 2940
aagccctccc gtatcgtagt tatctacacg acggggagtc aggcaactat ggatgaacga 3000
aatagacaga tcgctgagat aggtgcctca ctgattaagc attggtaact gtcagaccaa 3060
gtttactcat atatacttta gattgattta aaacttcatt tttaatttaa aaggatctag 3120
gtgaagatcc tttttgataa tctcatgacc aaaatccctt aacgtgagtt ttcgttccac 3180
tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt ttttctgcgc 3240
gtaatctgct gcttgcaaac aaaaaaacca ccgctaccag cggtggtttg tttgccggat 3300
caagagctac caactctttt tccgaaggta actggcttca gcagagcgca gataccaaat 3360
actgttcttc tagtgtagcc gtagttaggc caccacttca agaactctgt agcaccgcct 3420
acatacctcg ctctgctaat cctgttacca gtggctgctg ccagtggcga taagtcgtgt 3480
cttaccgggt tggactcaag acgatagtta ccggataagg cgcagcggtc gggctgaacg 3540
gggggttcgt gcacacagcc cagcttggag cgaacgacct acaccgaact gagataccta 3600
cagcgtgagc tatgagaaag cgccacgctt cccgaaggga gaaaggcgga caggtatccg 3660

gtaagcggca gggtcggaac aggagagcgc acgagggagc ttccaggggg aaacgcctgg 3720
tatctttata gtcctgtcgg gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc 3780
tcgtcagggg ggcggagcct atggaaaaac gccagcaacg cggccttttt acggttcctg 3840
gccttttgct ggccttttgc tcacat 3866

<210> 261

<211> 5618
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15776

<400> 261
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggtg gatctacgca cgctcggcta cagccagcag 480
caacaggaga agatcaaacc gaaggttcgt tcgacagtgg cgcagcacca cgaggcactg 540
gtcggccacg ggtttacaca cgcgcacatc gttgcgttaa gccaacaccc ggcagcgtta 600
gggaccgtcg ctgtcaagta tcaggacatg atcgcagcgt tgccagaggc gacacacgaa 660
gcgatcgttg gcgtcggcaa acagtggtcc ggcgcacgcg ctctggaggc cttgctcacg 720
gtggcgggag agttgagagg tccaccgtta cagttggaca caggccaact tctcaagatt 780
gcaaaacgtg gcggcgtgac cgcagtggag gcagtgcatg catggcgcaa tgcactgacg 840
ggtgccccgc tcaacttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 900
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 960
acccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 1020
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 1080
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 1140
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 1200
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 1260
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg 1320
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag 1380
gtggtggcca tcgccagcaa tattggtggc aagcaggcgc tggagacggt gcaggcgctg 1440
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 1500
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 1560
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag 1620
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 1680
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 1740
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 1800
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 1860
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 1920
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 1980
accccggagc aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg 2040
gtgcaggcgc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 2100
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 2160
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat 2220
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 2280
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg 2340
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 2400
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 2460
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 2520
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2580
gcccacggct tgacccctca gcaggtggtg gccatcgcca gcaatggcgg cggcaggccg 2640
gcgctggaga gcattgttgc ccagttatct cgccctgatc cgggatccaa gtctggcatc 2700
taccagatta agaacaccct gaataacaaa gtgtatgttg gatctgccaa ggactttgag 2760
aaaagatgga agaggcactt caaggatctg gaaaagggtt gccacagctc catcaagctc 2820
caaaggtctt tcaacaaaca tggcaatgtc tttgagtgct ccattttgga agagatccct 2880
tatgaaaagg acctcattat agagagggaa aacttctgga tcaaggagct gaactcaaaa 2940
atcaatggat acaacattgc agatgccaca tttggtgaca cctgttctac tcacccccctt 3000
aaggaagaga tcattaaaaa gaggtctgag actgtgaagg ctaagatgtt gaaacaaggt 3060
ccatcggccg actgataagt cgaccaccac caccaccacc actgataaga gctcctcgag 3120

cctgcagcag ctgaagcttt ggacttctgg cgcgcccaat tcgccctata gtgagtcgta 3180
ttacgtcgcg ctcactggcc gtcgttttac aacgtcgtga ctgggaaaac cctggcgtta 3240
cccaacttaa tcgccttgca gcacatcccc ctttcgccag ctggcgtaat agcgaagagg 3300
cccgcaccga aacgcccttc ccaacagttg cgcagcctga atggcgaatg ggagcgccct 3360
gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc gctacacttg 3420
ccagcgccct agcgcccgct cctttcgctt tcttcccttc ctttctcgcc acgttcgccg 3480
gctttccccg tcaagctcta aatcgggggc tcccctttagg gttccgattt agtgctttac 3540
ggcacctcga ccccaaaaaa cttgattagg gtgatggttg gcctgtagtg ggccatagcc 3600
ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg ttctttaata gtggactctt 3660
gttccaaact ggaacaacac tcaaccctat ctcggtctat tcttttgatt tataagggat 3720
tttgccgatt tcggcctatt ggttaaaaaa tgagctgatt taacaaaaat ttaacgcgaa 3780
ttttaacaaa atattaacgc ttacaattta ggtggcactt ttcggggaaa tgtgcgcgga 3840
acccctattt gtttattttt ctaaatacat tcaaatatgt atccgctcat gagacaataa 3900
ccctgataaa tgcttcaata atattgaaaa aggaagagta tgagtattca acatttccgt 3960
gtcgccctta ttcccttttt tgcggcattt tgccttcctg tttttgctca cccagaaacg 4020
ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac gagtgggtta catcgaactg 4080
gatctcaaca gcggtaagat ccttgagagt tttcgccccg aagaacgttt tccaatgatg 4140
agcacttta aagttctgct atgtggcgcg gtattatccc gtattgacgc cgggcaagag 4200
caactcggtc gccgcataca ctattctcag aatgacttgg ttgagtactc accagtcaca 4260
gaaaagcatc ttacggatgg catgacagta agagaattat gcagtgctgc cataaccatg 4320
agtgataaca ctgcggccaa cttacttctg acaacgatcg gaggaccgaa ggagctaacc 4380
gctttttgc acaacatggg ggatcatgta actcgccttg atcgttggga accggagctg 4440
aatgaagcca taccaaacga cgagcgtgac accacgatgc ctgtagcaat ggcaacaacg 4500
ttgcgcaaac tattaactgg cgaactactt actctagctt cccggcaaca attaatagac 4560
tggatggagg cggataaagt tgcaggacca cttctgcgct cggcccttcc ggctggctgg 4620
tttattgctg ataaatctgg agccggtgag cgtggttctc gcggtatcat tgcagcactg 4680
gggccagatg gtaagccctc ccgtatcgta gttatctaca cgacggggag tcaggcaact 4740
atggatgaac gaaatagaca gatcgctgag ataggtgcct cactgattaa gcattggtaa 4800
ctgtcagacc aagtttactc atatatactt tagattgatt taaaacttca tttttaattt 4860
aaaaggatct aggtgaagat cctttttgat aatctcatga ccaaaatccc ttaacgtgag 4920
ttttcgttcc actgagcgtc agaccccgta gaaaagatca aaggatcttc ttgagatcct 4980
ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc agcggtggtt 5040
tgtttgccgg atcaagagct accaactctt tttccgaagg taactggctt cagcagagcg 5100
cagataccaa atactgttct tctagtgtag ccgtagttag gccaccactt caagaactct 5160
gtagcaccgc ctacatacct cgctctgcta atcctgttac cagtggctgc tgccagtggc 5220
gataagtcgt gtcttaccgg gttggactca agacgatagt taccggataa ggcgcagcgg 5280
tcgggctgaa cgggggttc gtgcacacag cccagcttgg agcgaacgac ctacaccgaa 5340
ctgagatacc tacagcgtga gctatgagaa agcgccacgc ttcccgaagg gagaaaggcg 5400
gacaggtatc cggtaagcgg cagggtcgga acaggagagc gcacgaggga gcttccaggg 5460
ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact tgagcgtcga 5520
tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa acgccagcaa cgcggccttt 5580
ttacggttcc tggccttttg ctggcctttt gctcacat 5618

<210> 262

<211> 10929
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS12092

<400> 262
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
tttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840

gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg tggatctacg cacgctcggc 2460
tacagccagc agcaacagga gaagatcaaa ccgaaggttc gttcgacagt ggcgcagcac 2520
cacgaggcac tggtcggcca cgggtttaca cacgcgcaca tcgttgcgtt aagccaacac 2580
ccggcagcgt tagggaccgt cgctgtcaag tatcaggaca tgatcgcagc gttgccagag 2640
gcgacacacg aagcgatcgt tggcgtcggc aaacagtggt ccggcgcacg cgctctggag 2700
gccttgctca cggtggcggg agagttgaga ggtccaccgt tacagttgga cacaggccaa 2760
cttctcaaga ttgcaaaacg tggcggcgtg accgcagtgg aggcagtgca tgcatggcgc 2820
aatgcactga cgggtgcccc gctcaacttg accccggagc aggtggtggc catcgccagc 2880
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2940
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 3000
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 3060
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 3120
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 3180
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3240
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 3300
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 3360
accccggagc aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg 3420
gtgcaggcgc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 3480
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 3540
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat 3600
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 3660
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg 3720
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 3780
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 3840
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 3900
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 3960
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag 4020
gcgctggaga cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 4080
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 4140
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 4200
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 4260
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 4320
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 4380
acccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 4440
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 4500
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 4560
gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg tggccatcgc cagcaatggc 4620
ggcggcaggc cggcgctgga gagcattgtt gcccagttat ctcgccctga tccgggatcc 4680
aagtctggca tctaccagat taagaacacc ctgaataaca aagtgtatgt tggatctgcc 4740
aaggactttg agaaaagatg gaagaggcac ttcaaggatc tggaaaaggg ttgccacagc 4800
tccatcaagc tccaaaggtc tttcaacaaa catggcaatg tctttgagtg ctccattttg 4860
gaagagatcc cttatgaaaa ggacctcatt atagagaggg aaaacttctg gatcaaggag 4920

ctgaactcaa aaatcaatgg atacaacatt gcagatgcca catttggtga cacctgttct 4980
actcaccccc ttaaggaaga gatcattaaa aagaggtctg agactgtgaa ggctaagatg 5040
ttgaaacaag gtccatcggc cgactgataa ctcgagcgat cctctagacg agctcctcga 5100
gcctgcagca gctgaagctt tggacttctt cgccagaggt ttggtcaagt ctccaatcaa 5160
ggttgtcggc ttgtctacct tgccagaaat ttacgaaaag atggaaaagg gtcaaatcgt 5220
tggtagatac gttgttgaca cttctaaata agcgaatttc ttatgattta tgatttttat 5280
tattaaataa gttataaaaa aaataagtgt atacaaattt taaagtgact cttaggtttt 5340
aaaacgaaaa ttcttattct tgagtaactc tttcctgtag gtcaggttgc tttctcaggt 5400
atagcatgag gtcgctctta ttgaccacac ctctaccggc atgcaagctt ggcgtaatca 5460
tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga 5520
gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact cacattaatt 5580
gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagca gatctattac 5640
attatgggtg gtatgttgga ataaaaatca actatcatct actaactagt atttacgtta 5700
ctagtatatt atcatatacg gtgttagaag atgacgcaaa tgatgagaaa tagtcatcta 5760
aattagtgga agctgaaacg caaggattga taatgtaata ggatcaatga atattaacat 5820
ataaaatgat gataataata tttatagaat tgtgtagaat tgcagattcc cttttatgga 5880
ttcctaaatc ctcgaggaga acttctagta tatctacata cctaatatta ttgccttatt 5940
aaaaatggaa tcccaacaat tacatcaaaa tccacattct cttcaaaatc aattgtcctg 6000
tacttccttg ttcatgtgtg ttcaaaaacg ttatatttat aggataatta tactctattt 6060
ctcaacaagt aattggttgt ttggccgagc ggtctaaggc gcctgattca agaaatatct 6120
tgaccgcagt taactgtggg aatactcagg tatcgtaaga tgcaagagtt cgaatctctt 6180
agcaaccatt attttttttcc tcaacataac gagaacacac aggggcgcta tcgcacagaa 6240
tcaaattcga tgactggaaa ttttttgtta atttcagagg tcgcctgacg catatacctt 6300
tttcaactga aaaattggga gaaaaaggaa aggtgagagc cgcggaaccg gcttttcata 6360
tagaatagag aagcgttcat gactaaatgc ttgcatcaca atacttgaag ttgacaatat 6420
tatttaagga cctattgttt tttccaatag gtggttagca atcgtcttac tttctaactt 6480
ttcttacctt ttacatttca gcaatatata tatatatatt tcaaggatat accattctaa 6540
tgtctgcccc taagaagatc gtcgttttgc caggtgacca cgttggtcaa gaaatcacag 6600
ccgaagccat taaggttctt aaagctattt ctgatgttcg ttccaatgtc aagttcgatt 6660
tcgaaaatca tttaattggt ggtgctgcta tcgatgctac aggtgtccca cttccagatg 6720
aggcgctgga agcctccaag aaggttgatg ccgttttgtt aggtgctgtg ggtggtccta 6780
aatggggtac cggtagtgtt agacctgaac aaggtttact aaaaatccgt aaagaacttc 6840
aattgtacgc caacttaaga ccatgtaact ttgcatccga ctctctttta gacttatctc 6900
caatcaagcc acaatttgct aaaggtactg acttcgttgt tgtcagagaa ttagtgggag 6960
gtatttactt tggtaagaga aaggaagacg atggtgatgg tgtcgcttgg gatagtgaac 7020
aatacaccgt tccagaagtg caaagaatca caagaatggc cgctttcatg gccctacaac 7080
atgagccacc attgcctatt tggtccttgg ataaagctaa tgttttggcc tcttcaagat 7140
tatggagaaa aactgtggag gaaaccatca agaacgaatt ccctacattg aaggttcaac 7200
atcaattgat tgattctgcc gccatgatcc tagttaagaa cccaacccac ctaaatggta 7260
ttataatcac cagcaacatg tttggtgata tcatctccga tgaagcctcc gttatcccag 7320
gttccttggg tttgttgcca tctgcgtcct tggcctcttt gccagacaag aacaccgcat 7380
ttggtttgta cgaaccatgc cacggttctg ctccagattt gccaaagaat aaggtcaacc 7440
ctatcgccac tatcttgtct gctgcaatga tgttgaaatt gtcattgaac ttgcctgaag 7500
aaggtaaggc cattgaagat gcagttaaaa aggttttgga tgcaggtatc agaactggtg 7560
atttaggtgg ttccaacagt accacggaag tcggtgatgc tgtcgccgaa gaagttaaga 7620
aaatccttgc ttaaaaagat tctctttttt tatgatattt gtacataaac tttataaatg 7680
aaattcataa tagaaacgac acgaaattac aaaatggaat atgttcatag ggtagacgaa 7740
actatatacg caatctacat acatttatca agaaggagaa aaaggaggat gtaaaggaat 7800
acaggtaagc aaattgatac taatggctca acgtgataag gaaaaagaat tgcactttaa 7860
cattaatatt gacaaggagg agggcaccac acaaaaagtt aggtgtaaca gaaaatcatg 7920
aaactatgat tcctaattta tatattggag gattttctct aaaaaaaaaa aaatacaaca 7980
aataaaaaac actcaatgac ctgaccattt gatggagttt aagtcaatac cttcttgaac 8040
catttcccat aatggtgaaa gttccctcaa gaattttact ctgtcagaaa cggccttaac 8100
gacgtagtcg acctcctctt cagtactaaa tctaccaata ccaaatctga tggaagaatg 8160
ggctaatgca tcatccttac ccagcgcatg taaaacataa gaaggttcta gggaagcaga 8220
tgtacaggct gaacccgagg ataatgcgat atcccttagt gccatcaata aagattctcc 8280
ttccacgtag gcgaaagaaa cgttaacaca ccctggataa cgatgatctg gagatccgtt 8340
caacgtggta tgttcagcgg ataatagacc tttgactaat ttatcggata gtcttttgat 8400
gtgagcttgg tcgttgtcaa attctttctt catcaatctc gcagcttcac caaatcccgc 8460
taccaatggg ggggccaaag taccagatct gctgcattaa tgaatcggcc aacgcgcggg 8520
gagaggcggt ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc 8580
ggtcgttcgg ctgcggcgag cggtatcagc atcgatgaat tccacggact atagactata 8640
ctagtatact ccgtctactg tacgatacac ttccgctcag gtccttgtcc tttaacgagg 8700
ccttaccact cttttgttac tctattgatc cagctcagca aaggcagtgt gatctaagat 8760
tctatcttcg cgatgtagta aaactagcta gaccgagaaa gagactagaa atgcaaaagg 8820
cacttctaca atggctgcca tcattattat ccgatgtgac gctgcagctt ctcaatgata 8880
ttcgaatacg ctttgaggag atacagccta atatccgaca aactgtttta cagatttacg 8940
atcgtacttg ttacccatca ttgaattttg aacatccgaa cctgggagtt ttccctgaaa 9000

cagatagtat atttgaacct gtataataat atatagtcta gcgctttacg gaagacaatg 9060
tatgtatttc ggttcctgga gaaactattg catctattgc ataggtaatc ttgcacgtcg 9120
catccccggt tcatttttctg cgtttccatc ttgcacttca atagcatatc tttgttaacg 9180
aagcatctgt gcttcatttt gtagaacaaa aatgcaacgc gagagcgcta atttttcaaa 9240
caaagaatct gagctgcatt tttacagaac agaaatgcaa cgcgaaagcg ctattttacc 9300
aacgaagaat ctgtgcttca tttttgtaaa acaaaaatgc aacgcgagag cgctaatttt 9360
tcaaacaaag aatctgagct gcatttttac agaacagaaa tgcaacgcga gagcgctatt 9420
ttaccaacaa agaatctata cttctttttt gttctacaaa aatgcatccc gagagcgcta 9480
tttttctaac aaagcatctt agattacttt ttttctcctt tgtgcgctct ataatgcagt 9540
ctcttgataa ctttttgcac tgtaggtccg ttaaggttag aagaaggcta ctttggtgtc 9600
tattttctct tccataaaaa aagcctgact ccacttcccg cgtttactga ttactagcga 9660
agctgcgggt gcatttttttc aagataaagg catccccgat tatattctat accgatgtgg 9720
attgcgcata ctttgtgaac agaaagtgat agcgttgatg attcttcatt ggtcagaaaa 9780
ttatgaacgg tttcttctat tttgtctcta tatactacgt ataggaaatg tttacatttt 9840
cgtattgttt tcgattcact ctatgaatag ttcttactac aatttttttg tctaaagagt 9900
aatactagag ataaacataa aaaatgtaga ggtcgagttt agatgcaagt tcaaggagcg 9960
aaaggtggat gggtaggtta tatagggata tagcacagag atatatagca aagagatact 10020
tttgagcaat gtttgtggaa gcggtattcg caatatttta gtagctcgtt acagtccggt 10080
gcgtttttgg ttttttgaaa gtgcgtcttc agagcgcttt tggttttcaa aagcgctctg 10140
aagttcctat actttctaga gaataggaac ttcggaatag gaacttcaaa gcgtttccga 10200
aaacgagcgc ttccgaaaat gcaacgcgag ctgcgcacat acagctcact gttcacgtcg 10260
cacctatatc tgcgtgttgc ctgtatatat atatacatga gaagaacggc atagtgcgtg 10320
tttatgctta aatgcgtact tatatgcgtc tatttatgta ggatgaaagg tagtctagta 10380
cctcctgtga tattatccca ttccatgcgg ggtatcgtat gcttccttca gcactaccct 10440
ttagctgttc tatatgctgc cactcctcaa ttggattagt ctcatccttc aatgctatca 10500
tttcctttga tattggatca tatgcatagt accgagaaac tagtgcgaag tagtgatcag 10560
gtattgctgt tatctgatga gtatacgttg tcctggccac ggcagaagca cgcttatcgc 10620
tccaatttcc cacaacatta gtcaactccg ttaggccctt cattgaaaga aatgaggtca 10680
tcaaatgtct tccaatgtga gattttgggc catttttat agcaaagatt gaataaggcg 10740
cattttcctt caaagcttta ttgtacgatc tgactaagtt atcttttaat aattggtatt 10800
cctgtttatt gcttgaagaa ttgccggtcc tatttactcg ttttaggact ggttcagaat 10860
tcatcgatgc tcactcaaag gtcggtaata cggttatcca cagaatcagg ggataacgca 10920
ggaaagaac 10929

<210> 263

<211> 36
<212> DNA
<213> Artificial sequence

<220>
<223> CMP_G067

<400> 263
ggtatttccg gatggacctt gacctttcat cttgcc 36

<210> 264

<211> 53
<212> DNA
<213> Artificial sequence

<220>
<223> CMP_G152

<400> 264
atattttccg gagttagatg gtttcttacc tttcatcttg cctgaaattt tgg 53

<210> 265

<211> 36
<212> DNA
<213> Artificial sequence

<220>

<223> CMP_G114

<400> 265
ggtatttccg gatggacctt gtttcaacat cttagc 36

<210> 266

<211> 4094
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15777

<400> 266
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggcc aagtctggca tctaccagat taagaacacc 480
ctgaataaca aagtgtatgt tggatctgcc aaggactttg agaaaagatg gaagaggcac 540
ttcaaggatc tggaaaaggg ttgccacagc tccatcaagc tccaaaggtc tttcaacaaa 600
catggcaatg tctttgagtg ctccattttg gaagagatcc cttatgaaaa ggacctcatt 660
atagagaggg aaaacttctg gatcaaggag ctgaactcaa aaatcaatgg atacaacatt 720
gcagatgcca catttggtga cacctgttct actcaccccc ttaaggaaga gatcattaaa 780
aagaggtctg agactgtgaa ggctaagatg ttgaaactgg ggcctgatgg cagaaaggca 840
ctctactcta agccaggtag caaaaatgga aggtggaacc ctgagaccca caagttctgc 900
aaatgtgggg tgagaataca gacatcagcc tacacttgca gcaagtgtag gaacagatca 960
ggagagaata actccttttt caaccacaag cattctgaca tcaccaagtc caaaatttca 1020
ggcaagatga aaggtaagaa accatctaac tccggagtgg atctacgcac gctcggctac 1080
agccagcagc aacaggagaa gatcaaaccg aaggttcgtt cgacagtggc gcagcaccac 1140
gaggcactgg tcggccacgg gtttacacac gcgcacatcg ttgcgttaag ccaacacccg 1200
gcagcgttag ggaccgtcgc tgtcaagtat caggacatga tcgcagcgtt gccagaggcg 1260
acacacgaag cgatcgttgg cgtcggcaaa cagtggtccg gcgcacgcgc tctggaggcc 1320
ttgctcacgg tggcgggaga gttgagaggt ccaccgttac agttggacac aggccaactt 1380
ctcaagattg caaaacgtgg cggcgtgacc gcagtggagg cagtgcatgc atggcgcaat 1440
gcactgacgg gtgccccgct caacttgacc ggagacgccc gggggatcag gtcacgtgcg 1500
tctcggagca ttgttgccca gttatctcgc cctgatccgt cggccgactg ataagtcgac 1560
caccaccacc accaccactg ataagagctc ctcgagcctg cagcagctga agctttggac 1620
ttctggcgcg cccaattcgc cctatagtga gtcgtattac gtcgcgctca ctggccgtcg 1680
ttttacaacg tcgtgactgg gaaaaccctg gcgttaccca acttaatcgc cttgcagcac 1740
atccccctttt cgccagctgg cgtaatagcg aagaggcccg caccgaaacg cccttcccaa 1800
cagttgcgca gcctgaatgg cgaatgggag cgccctgtag cggcgcatta agcgcggcgg 1860
gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt 1920
tcgctttctt cccttccttt ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc 1980
gggggctccc tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg 2040
attagggtga tggttggcct gtagtgggcc atagccctga tagacggttt ttcgcccttt 2100
gacgttggag tccacgttct ttaatagtgg actcttgttc caaactggaa caacactcaa 2160
ccctatctcg gtctattctt ttgatttata agggattttg ccgatttcgg cctattggtt 2220
aaaaaatgag ctgatttaac aaaaatttaa cgcgaatttt aacaaaatat taacgcttac 2280
aatttaggtg gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt atttttctaa 2340
atacattcaa atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat 2400
tgaaaaagga agagtatgag tattcaacat ttccgtgtcg cccttattcc ctttttgcg 2460
gcattttgcc ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa 2520
gatcagttgg gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt 2580
gagagttttc gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt 2640
ggcgcggtat tatcccgtat tgacgccggg caagagcaac tcggtcgccg catacactat 2700
tctcagaatg acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg 2760
acagtaagag aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta 2820
cttctgacaa cgatcggagg accgaaggag ctaaccgctt ttttgcacaa catgggggat 2880
catgtaactc gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag 2940
cgtgacacca cgatgcctgt agcaatggca acaacgttgc gcaaactatt aactggcgaa 3000
ctacttactc tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca 3060
ggaccacttc tgcgctcggc ccttccggct ggctggttta ttgctgataa atctggagcc 3120

ggtgagcgtg gttctcgcgg tatcattgca gcactggggc cagatggtaa gccctcccgt 3180
atcgtagtta tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc 3240
gctgagatag gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat 3300
atactttaga ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt 3360
tttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac 3420
cccgtagaaa agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc 3480
ttgcaaacaa aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca 3540
actcttttc cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta 3600
gtgtagccgt agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct 3660
ctgctaatcc tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg 3720
gactcaagac gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc 3780
acacagccca gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta 3840
tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg 3900
gtcggaacag gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt 3960
cctgtcgggt ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcaggggg 4020
cggagcctat ggaaaaacgc cagcaacgcg gccttttac ggttcctggc cttttgctgg 4080
ccttttgctc acat 4094

<210> 267

<211> 4088
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15778

<400> 267
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggcc aagtctggca tctaccagat taagaacacc 480
ctgaataaca aagtgtatgt tggatctgcc aaggactttg agaaaagatg gaagaggcac 540
ttcaaggatc tggaaaaggg ttgccacagc tccatcaagc tccaaaggtc tttcaacaaa 600
catggcaatg tctttgagtg ctccattttg gaagagatcc cttatgaaaa ggacctcatt 660
atagagaggg aaaacttctg gatcaaggag ctgaactcaa aaatcaatgg atacaacatt 720
gcagatgcca catttggtga cacctgttct actcaccccc ttaaggaaga gatcattaaa 780
aagaggtctg agactgtgaa ggctaagatg ttgaaactgg ggcctgatgg cagaaaggca 840
ctctactcta agccaggtag caaaaatgga aggtggaacc ctgagaccca caagttctgc 900
aaatgtgggg tgagaataca gacatcagcc tacacttgca gcaagtgtag gaacagatca 960
ggagagaata actccttttt caaccacaag cattctgaca tcaccaagtc caaaatttca 1020
ggcaagatga aaggtcaagg tccatccgga gtggatctac gcacgctcgg ctacagccag 1080
cagcaacagg agaagatcaa accgaaggtt cgttcgacag tggcgcagca ccacgaggca 1140
ctggtcggcc acgggtttac acacgcgcac atcgttgcgt taagccaaca cccggcagcg 1200
ttagggaccg tcgctgtcaa gtatcaggac atgatcgcag cgttgccaga ggcgacacac 1260
gaagcgatcg ttggcgtcgg caaacagtgg tccggcgcac gcgctctgga ggccttgctc 1320
acggtggcgg gagagttgag aggtccaccg ttacagttgg acacaggcca acttctcaag 1380
attgcaaaac gtggcggcgt gaccgcagtg gaggcagtgc atgcatggcg caatgcactg 1440
acgggtgccc cgctcaactt gaccggagac gcccggggga tcaggtcacg tgcgtctcgg 1500
agcattgttg cccagttatc tcgccctgat ccgtcggccg actgataagt cgaccaccac 1560
caccaccacc actgataaga gctcctcgag cctgcagcag ctgaagcttt ggacttctgg 1620
cgcgcccaat tcgccctata gtgagtcgta ttacgtcgcg ctcactggcc gtcgttttac 1680
aacgtcgtga ctgggaaaac cctggcgtta cccaacttaa tcgccttgca gcacatcccc 1740
ctttcgccag ctggcgtaat agcgaagagg cccgcaccga aacgcccttc ccaacagttg 1800
cgcagcctga atggcgaatg ggagcgccct gtagcggcgc attaagcgcg gcgggtgtgg 1860
tggttacgcg cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt 1920
tcttcccttc ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc 1980
tcccctttagg gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg 2040
gtgatggttg gcctgtagtg ggccatagcc ctgatagacg gtttttcgcc ctttgacgtt 2100
ggagtccacg ttctttaata gtggactctt gttccaaact ggaacaacac tcaaccctat 2160
ctcggtctat tcttttgatt tataagggat tttgccgatt tcggcctatt ggttaaaaaa 2220
tgagctgatt taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc ttacaattta 2280
ggtggcactt ttcggggaaa tgtgcgcgga acccctattt gtttatttttt ctaaatacat 2340

tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa 2400
aggaagagta tgagtattca acatttccgt gtcgccctta ttccctttt tgcggcattt 2460
tgccttcctg tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag 2520
ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaagat ccttgagagt 2580
tttcgccccg aagaacgttt tccaatgatg agcacttta aagttctgct atgtggcgcg 2640
gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag 2700
aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta 2760
agagaattat gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg 2820
acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta 2880
actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac 2940
accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt 3000
actctagctt cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca 3060
cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag 3120
cgtggttctc gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta 3180
gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag 3240
ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt 3300
tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat 3360
aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta 3420
gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa 3480
acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt 3540
tttccgaagg taactggctt cagcagagcg cagataccaa atactgttct tctagtgtag 3600
ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta 3660
atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca 3720
agacgatagt taccggataa ggcgcagcgg tcgggctgaa cgggggttc gtgcacacag 3780
cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa 3840
agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga 3900
acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc 3960
gggtttcgcc acctctgact tgagcgtcga ttttgtgat gctcgtcagg ggggcggagc 4020
ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt 4080
gctcacat 4088

<210> 268

<211> 3869
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15779

<400> 268
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggcc aagtctggca tctaccagat taagaacacc 480
ctgaataaca aagtgtatgt tggatctgcc aaggactttg agaaaagatg gaagaggcac 540
ttcaaggatc tggaaaaggg ttgccacagc tccatcaagc tccaaaggtc tttcaacaaa 600
catggcaatg tctttgagtg ctccattttg gaagagatcc cttatgaaaa ggacctcatt 660
atagagaggg aaaacttctg gatcaaggag ctgaactcaa aaatcaatgg atacaacatt 720
gcagatgcca catttggtga cacctgttct actcaccccc ttaaggaaga gatcattaaa 780
aagaggtctg agactgtgaa ggctaagatg ttgaaacaag gtccatccgg agtggatcta 840
cgcacgctcg gctacagcca gcagcaacag gagaagatca aaccgaaggt tcgttcgaca 900
gtggcgcagc accacgaggc actggtcggc cacgggttta cacacgcgca catcgttgcg 960
ttaagccaac acccggcagc gttagggacc gtcgctgtca agtatcagga catgatcgca 1020
gcgttgccag aggcgacaca cgaagcgatc gttggcgtcg gcaaacagtg gtccggcgca 1080
cgcgctctgg aggccttgct cacggtggcg ggagagttga gaggtccacc gttacagttg 1140
gacacaggcc aacttctcaa gattgcaaaa cgtggcggcg tgaccgcagt ggaggcagtg 1200
catgcatggc gcaatgcact gacgggtgcc ccgctcaact tgaccggaga cgcccggggg 1260
atcaggtcac gtgcgtctcg gagcattgtt gcccagttat ctcgccctga tccgtcggcc 1320
gactgataag tcgaccacca ccaccaccac cactgataag agctcctcga gcctgcagca 1380
gctgaagctt tggacttctg gcgcgcccaa ttcgccctat agtgagtcgt attacgtcgc 1440
gctcactggc cgtcgtttta caacgtcgtg actgggaaaa ccctggcgtt acccaactta 1500
atcgccttgc agcacatccc cctttcgcca gctggcgtaa tagcgaagag gcccgcaccg 1560

aaacgccctt cccaacagtt gcgcagcctg aatggcgaat gggagcgccc tgtagcggcg 1620
cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgccc 1680
tagcgcccgc tcctttcgct ttcttccctt cctttctcgc cacgttcgcc ggctttcccc 1740
gtcaagctct aaatcggggg ctccctttag ggttccgatt tagtgcttta cggcacctcg 1800
accccaaaaa acttgattag ggtgatggtt ggcctgtagt gggccatagc cctgatagac 1860
ggtttttcgc cctttgacgt tggagtccac gttctttaat agtggactct tgttccaaac 1920
tggaacaaca ctcaacccta tctcggtcta ttcttttgat ttataaggga ttttgccgat 1980
ttcggcctat tggttaaaaa atgagctgat ttaacaaaaa tttaacgcga attttaacaa 2040
aatattaacg cttacaattt aggtggcact tttcggggaa atgtgcgcgg aacccctatt 2100
tgtttatttt tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa 2160
atgcttcaat aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt 2220
attccctttt ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa 2280
gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac 2340
agcggtaaga tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcacttt 2400
aaagttctgc tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggt 2460
cgccgcatac actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat 2520
cttacggatg gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac 2580
actgcggcca acttacttct gacaacgatc ggaggaccga aggagctaac cgcttttttg 2640
cacaacatgg gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc 2700
ataccaaacg acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa 2760
ctattaactg gcgaactact tactctagct tcccggcaac aattaataga ctggatggag 2820
gcggataaag ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct 2880
gataaatctg gagccggtga gcgtggttct cgcggtatca ttgcagcact ggggccagat 2940
ggtaagccct cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa 3000
cgaaatagac agatcgctga gataggtgcc tcactgatta agcattggta actgtcagac 3060
caagtttact catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc 3120
taggtgaaga tccttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc 3180
cactgagcgt cagaccccgt agaaaagatc aaaggatctt cttgagatcc ttttttctg 3240
cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg 3300
gatcaagagc taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca 3360
aatactgttc ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg 3420
cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg 3480
tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga 3540
acgggggttt cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac 3600
ctacagcgtg agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat 3660
ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc 3720
tggtatcttt atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga 3780
tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc 3840
ctggccttttt gctggccttt tgctcacat 3869

<210> 269

<211> 5846
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15780

<400> 269
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggcc aagtctggca tctaccagat taagaacacc 480
ctgaataaca aagtgtatgt tggatctgcc aaggactttg agaaaagatg gaagaggcac 540
ttcaaggatc tggaaaaggg ttgccacagc tccatcaagc tccaaaggtc tttcaacaaa 600
catggcaatg tctttgagtg ctccattttg gaagagatcc cttatgaaaa ggacctcatt 660
atagagaggg aaaacttctg gatcaaggag ctgaactcaa aaatcaatgg atacaacatt 720
gcagatgcca catttggtga cacctgttct actcaccccc ttaaggaaga gatcattaaa 780
aagaggtctg agactgtgaa ggctaagatg ttgaaactgg ggcctgatgg cagaaaggca 840
ctctactcta agccaggtag caaaaatgga aggtggaacc ctgagaccca caagttctgc 900
aaatgtgggg tgagaataca gacatcagcc tacacttgca gcaagtgtag gaacagatca 960
ggagagaata actccttttt caaccacaag cattctgaca tcaccaagtc caaaatttca 1020

ggcaagatga aaggtaagaa accatctaac tccggagtgg atctacgcac gctcggctac 1080
agccagcagc aacaggagaa gatcaaaccg aaggttcgtt cgacagtggc gcagcaccac 1140
gaggcactgg tcggccacgg gtttacacac gcgcacatcg ttgcgttaag ccaacacccg 1200
gcagcgttag ggaccgtcgc tgtcaagtat caggacatga tcgcagcgtt gccagaggcg 1260
acacacgaag cgatcgttgg cgtcggcaaa cagtggtccg gcgcacgcgc tctggaggcc 1320
ttgctcacgg tggcgggaga gttgagaggt ccaccgttac agttggacac aggccaactt 1380
ctcaagattg caaaacgtgg cggcgtgacc gcagtggagg cagtgcatgc atggcgcaat 1440
gcactgacgg gtgccccgct caacttgacc ccggagcagg tggtggccat cgccagccac 1500
gatggcggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc 1560
cacggcttga ccccccagca ggtggtggcc atcgccagca atggcggtgg caagcaggcg 1620
ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt gaccccggag 1680
caggtggtgg ccatcgccag caatattggt ggcaagcagg cgctggagac ggtgcaggcg 1740
ctgttgccgg tgctgtgcca ggcccacggc ttgacccccc agcaggtggt ggccatcgcc 1800
agcaatggcg gtggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc 1860
caggcccacg gcttgacccc ggagcaggtg gtggccatcg ccagcaatat tggtggcaag 1920
caggcgctgg agacggtgca ggcgctgttg ccggtgctgt gccaggccca cggcttgacc 1980
ccggagcagg tggtggccat cgccagcaat attggtggca agcaggcgct ggagacggtg 2040
caggcgctgt tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc 2100
atcgccagca atattggtgg caagcaggcg ctggagacgg tgcaggcgct gttgccggtg 2160
ctgtgccagg cccacggctt gaccccggag caggtggtgg ccatcgccag ccacgatggc 2220
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc 2280
ttgaccccgg agcaggtggt ggccatcgcc agccacgatg gcggcaagca ggcgctggag 2340
acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc ccagcaggtg 2400
gtggccatcg ccagcaatgg cggtggcaag caggcgctgg agacggtcca gcggctgttg 2460
ccggtgctgt gccaggccca cggcttgacc ccggagcagg tggtggccat cgccagcaat 2520
attggtggca agcaggcgct ggagacggtg caggcgctgt tgccggtgct gtgccaggcc 2580
cacggcttga ccccggagca ggtggtggcc atcgccagca atattggtgg caagcaggcg 2640
ctggagacgg tgcaggcgct gttgccggtg ctgtgccagg cccacggctt gaccccggag 2700
caggtggtgg ccatcgccag ccacgatggc ggcaagcagg cgctggagac ggtccagcgg 2760
ctgttgccgg tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc 2820
agccacgatg gcggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc 2880
caggcccacg gcttgacccc ggagcaggtg gtggccatcg ccagccacga tggcggcaag 2940
caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc 3000
ccccagcagg tggtggccat cgccagcaat ggcggtggca agcaggcgct ggagacggtc 3060
cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc 3120
atcgccagcc acgatggcgg caagcaggcg ctggagacgg tccagcggct gttgccggtg 3180
ctgtgccagg cccacggctt gacccctcag caggtggtgg ccatcgccag caatggcggc 3240
ggcaggccgg cgctggagag cattgttgcc cagttatctc gccctgatcc gtcggccgac 3300
tgataagtcg accaccacca ccaccaccac tgataagagc tcctcgagcc tgcagcagct 3360
gaagctttgg acttctggcg cgcccaattc gccctatagt gagtcgtatt acgtcgcgct 3420
cactggccgt cgttttacaa cgtcgtgact gggaaaaccc tggcgttacc caacttaatc 3480
gccttgcagc acatccccct ttcgccagct ggcgtaatag cgaagaggcc cgcaccgaaa 3540
cgcccttccc aacagttgcg cagcctgaat ggcgaatggg agcgccctgt agcggcgcat 3600
taagcgcggc gggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc agcgccctag 3660
cgcccgctcc tttcgctttc ttcccttcct ttctcgccac gttcgccggc tttccccgtc 3720
aagctctaaa tcgggggctc cctttagggt tccgatttag tgctttacgg cacctcgacc 3780
ccaaaaaact tgattagggt gatggttggc ctgtagtggg ccatagccct gatagacggt 3840
ttttcgccct ttgacgttgg agtccacgtt ctttaatagt ggactcttgt tccaaactgg 3900
aacaacactc aaccctatct cggtctattc ttttgattta taagggattt tgccgatttc 3960
ggcctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt ttaacaaaat 4020
attaacgctt acaatttagg tggcactttt cggggaaatg tgcgcggaac ccctatttgt 4080
ttatttttct aaatacattc aaatatgtat ccgctcatga gacaataacc ctgataaatg 4140
cttcaataat attgaaaaag gaagagtatg agtattcaac atttccgtgt cgcccttatt 4200
ccctttttg cggcattttg ccttcctgtt tttgctcacc cagaaacgct ggtgaaagta 4260
aaagatgctg aagatcagtt gggtgcacga gtgggttaca tcgaactgga tctcaacagc 4320
ggtaagatcc ttgagagttt tcgccccgaa gaacgttttc caatgatgag cacttttaaa 4380
gttctgctat gtggcgcggt attatcccgt attgacgccg ggcaagagca actcggtcgc 4440
cgcatacact attctcagaa tgacttggtt gagtactcac cagtcacaga aaagcatctt 4500
acggatggca tgacagtaag agaattatgc agtgctgcca taaccatgag tgataacact 4560
gcggccaact tacttctgac aacgatcgga ggaccgaagg agctaaccgc ttttttgcac 4620
aacatggggg atcatgtaac tcgccttgat cgttgggaac cggagctgaa tgaagccata 4680
ccaaacgacg agcgtgacac cacgatgcct gtagcaatgg caacaacgtt gcgcaaacta 4740
ttaactggcg aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg 4800
gataaagttg caggaccact tctgcgctcg gcccttccgg ctggctggtt tattgctgat 4860
aaatctggag ccggtgagcg tggttctcgc ggtatcattg cagcactggg gccagatggt 4920
aagccctccc gtatcgtagt tatctacacg acggggagtc aggcaactat ggatgaacga 4980
aatagacaga tcgctgagat aggtgcctca ctgattaagc attggtaact gtcagaccaa 5040
gtttactcat atatacttta gattgattta aaacttcatt tttaatttaa aaggatctag 5100

gtgaagatcc tttttgataa tctcatgacc aaaatccctt aacgtgagtt ttcgttccac 5160
tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt ttttctgcgc 5220
gtaatctgct gcttgcaaac aaaaaaacca ccgctaccag cggtggtttg tttgccggat 5280
caagagctac caactctttt tccgaaggta actggcttca gcagagcgca gataccaaat 5340
actgttcttc tagtgtagcc gtagttaggc caccacttca agaactctgt agcaccgcct 5400
acatacctcg ctctgctaat cctgttacca gtggctgctg ccagtggcga taagtcgtgt 5460
cttaccgggt tggactcaag acgatagtta ccggataagg cgcagcggtc gggctgaacg 5520
gggggttcgt gcacacagcc cagcttggag cgaacgacct acaccgaact gagataccta 5580
cagcgtgagc tatgagaaag cgccacgctt cccgaaggga gaaaggcgga caggtatccg 5640
gtaagcggca gggtcggaac aggagagcgc acgagggagc ttccaggggg aaacgcctgg 5700
tatctttata gtcctgtcgg gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc 5760
tcgtcagggg ggcggagcct atggaaaaac gccagcaacg cggccttttt acggttcctg 5820
gccttttgct ggccttttgc tcacat 5846

<210> 270

<211> 5840
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15781

<400> 270
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggcc aagtctggca tctaccagat taagaacacc 480
ctgaataaca aagtgtatgt tggatctgcc aaggactttg agaaaagatg gaagaggcac 540
ttcaaggatc tggaaaaggg ttgccacagc tccatcaagc tccaaaggtc tttcaacaaa 600
catggcaatg tctttgagtg ctccattttg gaagagatcc cttatgaaaa ggacctcatt 660
atagagaggg aaaacttctg gatcaaggag ctgaactcaa aaatcaatgg atacaacatt 720
gcagatgcca catttggtga cacctgttct actcaccccc ttaaggaaga gatcattaaa 780
aagaggtctg agactgtgaa ggctaagatg ttgaaactgg ggcctgatgg cagaaaggca 840
ctctactcta agccaggtag caaaaatgga aggtggaacc ctgagaccca caagttctgc 900
aaatgtgggg tgagaataca gacatcagcc tacacttgca gcaagtgtag gaacagatca 960
ggagagaata actccttttt caaccacaag cattctgaca tcaccaagtc caaaatttca 1020
ggcaagatga aaggtcaagg tccatccgga gtggatctac gcacgctcgg ctacagccag 1080
cagcaacagg agaagatcaa accgaaggtt cgttcgacag tggcgcagca ccacgaggca 1140
ctggtcggcc acgggtttac acacgcgcac atcgttgcgt taagccaaca cccggcagcg 1200
ttagggaccg tcgctgtcaa gtatcaggac atgatcgcag cgttgccaga ggcgacacac 1260
gaagcgatcg ttggcgtcgg caaacagtgg tccggcgcac gcgctctgga ggccttgctc 1320
acggtggcgg gagagttgag aggtccaccg ttacagttgg acacaggcca acttctcaag 1380
attgcaaaac gtggcggcgt gaccgcagtg gaggcagtgc atgcatggcg caatgcactg 1440
acgggtgccc cgctcaactt gaccccggag caggtggtgg ccatcgccag ccacgatggc 1500
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc 1560
ttgacccccc agcaggtggt ggccatcgcc agcaatggcg gtggcaagca ggcgctggag 1620
acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg 1680
gtggccatcg ccagcaatat tggtggcaag caggcgctgg agacggtgca ggcgctgttg 1740
ccggtgctgt gccaggccca cggcttgacc ccccagcagg tggtggccat cgccagcaat 1800
ggcggtggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc 1860
cacggcttga ccccggagca ggtggtggcc atcgccagca atattggtgg caagcaggcg 1920
ctggagacgg tgcaggcgct gttgccggtg ctgtgccagg cccacggctt gaccccggag 1980
caggtggtgg ccatcgccag caatattggt ggcaagcagg cgctggagac ggtgcaggcg 2040
ctgttgccgg tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc 2100
agcaatattg gtggcaagca ggcgctggag acggtgcagg cgctgttgcc ggtgctgtgc 2160
caggcccacg gcttgacccc ggagcaggtg gtggccatcg ccagccacga tggcggcaag 2220
caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc 2280
ccggagcagg tggtggccat cgccagccac gatggcggca agcaggcgct ggagacggtc 2340
cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccccagca ggtggtggcc 2400
atcgccagca atggcggtgg caagcaggcg ctggagacgg tccagcggct gttgccggtg 2460
ctgtgccagg cccacggctt gaccccggag caggtggtgg ccatcgccag caatattggt 2520
ggcaagcagg cgctggagac ggtgcaggcg ctgttgccgg tgctgtgcca ggcccacggc 2580

ttgaccccgg agcaggtggt ggccatcgcc agcaatattg gtggcaagca ggcgctggag 2640
acggtgcagg cgctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg 2700
gtggccatcg ccagccacga tggcggcaag caggcgctgg agacggtcca gcggctgttg 2760
ccggtgctgt gccaggccca cggcttgacc ccggagcagg tggtggccat cgccagccac 2820
gatggcggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc 2880
cacggcttga ccccggagca ggtggtggcc atcgccagcc acgatggcgg caagcaggcg 2940
ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt gacccccag 3000
caggtggtgg ccatcgccag caatggcggt ggcaagcagg cgctggagac ggtccagcgg 3060
ctgttgccgg tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc 3120
agccacgatg gcggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc 3180
caggcccacg gcttgacccc tcagcaggtg gtggccatcg ccagcaatgg cggcggcagg 3240
ccggcgctgg agagcattgt tgcccagtta tctcgccctg atccgtcggc cgactgataa 3300
gtcgaccacc accaccacca ccactgataa gagctcctcg agcctgcagc agctgaagct 3360
ttggacttct ggcgcgccca attcgcccta tagtgagtcg tattacgtcg cgctcactgg 3420
ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt tacccaactt aatcgccttg 3480
cagcacatcc ccctttcgcc agctggcgta atagcgaaga ggcccgcacc gaaacgccct 3540
tcccaacagt tgcgcagcct gaatggcgaa tgggagcgcc ctgtagcggc gcattaagcg 3600
cggcgggtgt ggtggttacg cgcagcgtga ccgctacact tgccagcgcc ctagcgcccg 3660
ctcctttcgc tttcttccct tcctttctcg ccacgttcgc cggctttccc cgtcaagctc 3720
taaatcgggg gctcccttta gggttccgat ttagtgcttt acggcacctc gaccccaaaa 3780
aacttgatta gggtgatggt tggcctgtag tgggccatag ccctgataga cggtttttcg 3840
ccctttgacg ttggagtcca cgttctttaa tagtggactc ttgttccaaa ctggaacaac 3900
actcaaccct atctcggtct attcttttga tttataaggg attttgccga tttcggccta 3960
ttggttaaaa aatgagctga tttaacaaaa atttaacgcg aattttaaca aaatattaac 4020
gcttacaatt taggtggcac ttttcgggga aatgtgcgcg gaacccctat ttgtttattt 4080
ttctaaatac attcaaatat gtatccgctc atgagacaat aaccctgata aatgcttcaa 4140
taatattgaa aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattcccttt 4200
tttgcggcat tttgccttcc tgtttttgct cacccagaaa cgctggtgaa agtaaaagat 4260
gctgaagatc agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag 4320
atccttgaga gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagttctg 4380
ctatgtggcg cggtattatc ccgtattgac gccgggcaag agcaactcgg tcgccgcata 4440
cactattctc agaatgactt ggttgagtac tcaccagtca cagaaaagca tcttacggat 4500
ggcatgacag taagagaatt atgcagtgct gccataacca tgagtgataa cactgcggcc 4560
aacttacttc tgacaacgat cggaggaccg aaggagctaa ccgcttttt gcacaacatg 4620
ggggatcatg taactcgcct tgatcgttgg gaaccggagc tgaatgaagc cataccaaac 4680
gacgagcgtg acaccacgat gcctgtagca atggcaacaa cgttgcgcaa actattaact 4740
ggcgaactac ttactctagc ttcccggcaa caattaatag actggatgga ggcggataaa 4800
gttgcaggac cacttctgcg ctcggccctt ccggctggct ggtttattgc tgataaatct 4860
ggagccggtg agcgtggttc tcgcggtatc attgcagcac tggggccaga tggtaagccc 4920
tcccgtatcg tagttatcta cacgacgggg agtcaggcaa ctatggatga acgaaataga 4980
cagatcgctg agataggtgc ctcactgatt aagcattggt aactgtcaga ccaagtttac 5040
tcatatatac tttagattga tttaaaactt catttttaat ttaaaaggat ctaggtgaag 5100
atccttttg ataatctcat gaccaaaatc ccttaacgtg agttttcgtt ccactgagcg 5160
tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc 5220
tgctgcttgc aaacaaaaaa accaccgcta ccagcggtgg tttgtttgcc ggatcaagag 5280
ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtt 5340
cttctagtgt agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac 5400
ctcgctctgc taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc 5460
gggttggact caagacgata gttaccggat aaggcgcagc ggtcgggctg aacgggggt 5520
tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt 5580
gagctatgag aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggtaagc 5640
ggcagggtcg gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt 5700
tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca 5760
ggggggcgga gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt 5820
tgctggcctt ttgctcacat 5840

<210> 271

<211> 1614
<212> DNA
<213> Artificial sequence

<220>
<223> RagT2-R RVD insert

<400> 271
gcagcctcaa cttgaccccc cagcaggtgg tggccatcgc cagcaataat ggtggcaagc 60

aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 120
cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc 180
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag gtggtggcca 240
tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 300
tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc aatggcggtg 360
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 420
tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag gcgctggaga 480
cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg 540
tggccatcgc cagcaatggc ggtggcaagc aggcgctgga gacggtccag cggctgttgc 600
cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc gccagcaata 660
atggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg tgccaggccc 720
acggcttgac cccccagcag gtggtggcca tcgccagcaa taatggtggc aagcaggcgc 780
tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg acccccccagc 840
aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg gtccagcggc 900
tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg gccatcgcca 960
gcaatggcgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc 1020
aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc 1080
aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc 1140
cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg gagacggtcc 1200
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag gtggtggcca 1260
tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 1320
tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc aatggcggtg 1380
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 1440
tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag gcgctggaga 1500
cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg 1560
tggccatcgc cagcaatggc ggcggcaggc cggcgctgga gagcattgtg atgc 1614

<210> 272

<211> 5417
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15782

<400> 272
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggcc aagtctggca tctaccagat taagaacacc 480
ctgaataaca aagtgtatgt tggatctgcc aaggactttg agaaaagatg gaagaggcac 540
ttcaaggatc tggaaaaggg ttgccacagc tccatcaagc tccaaaggtc tttcaacaaa 600
catggcaatg tctttgagtg ctccattttg gaagagatcc cttatgaaaa ggacctcatt 660
atagagaggg aaaacttctg gatcaaggag ctgaactcaa aaatcaatgg atacaacatt 720
gcagatgcca catttggtga cacctgttct actcaccccc ttaaggaaga gatcattaaa 780
aagaggtctg agactgtgaa ggctaagatg ttgaaacaag gtccatccgg agtggatcta 840
cgcacgctcg gctacagcca gcagcaacag gagaagatca aaccgaaggt tcgttcgaca 900
gtggcgcagc accacgaggc actggtcggc cacgggttta cacacgcgca catcgttgcg 960
ttaagccaac acccggcagc gttagggacc gtcgctgtca agtatcagga catgatcgca 1020
gcgttgccag aggcgacaca cgaagcgatc gttggcgtcg gcaaacagtg gtccggcgca 1080
cgcgctctgg aggccttgct cacggtggcg ggagagttga gaggtccacc gttacagttg 1140
gacacaggcc aacttctcaa gattgcaaaa cgtggcggcg tgaccgcagt ggaggcagtg 1200
catgcatggc gcaatgcact gacgggtgcc ccgctcaact tgaccccca gcaggtggtg 1260
gccatcgcca gcaataatgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg 1320
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 1380
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 1440
ggcttgaccc cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg 1500
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 1560
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 1620
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 1680
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 1740
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 1800

gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc 1860
cagcaggtgg tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag 1920
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 1980
gccagcaata atggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 2040
tgccaggccc acggcttgac cccccagcag gtggtggcca tcgccagcaa tggcggtggc 2100
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 2160
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 2220
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 2280
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 2340
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat 2400
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 2460
ggcttgaccc cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg 2520
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 2580
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 2640
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 2700
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 2760
gcccacggct tgacccctca gcaggtggtg gccatcgcca gcaatggcgg cggcaggccg 2820
gcgctggaga gcattgttgc ccagttatct cgccctgatc cgtcggccga ctgataagtc 2880
gaccaccacc accaccacca ctgataagag ctcctcgagc ctgcagcagc tgaagctttg 2940
gacttctggc gcgcccaatt cgccctatag tgagtcgtat tacgtcgcgc tcactggccg 3000
tcgttttaca acgtcgtgac tgggaaaacc ctggcgttac ccaacttaat cgccttgcag 3060
cacatccccc tttcgccagc tggcgtaata gcgaagaggc ccgcaccgaa acgcccttcc 3120
caacagttgc gcagcctgaa tggcgaatgg gagcgccctg tagcggcgca ttaagcgcgg 3180
cgggtgtggt ggttacgcgc agcgtgaccg ctacacttgc cagcgcccta gcgcccgctc 3240
ctttcgcttt cttcccttcc tttctcgcca cgttcgccgg ctttccccgt caagctctaa 3300
atcgggggct ccctttaggg ttccgattta gtgctttacg gcacctcgac cccaaaaaac 3360
ttgattaggg tgatggttgg cctgtagtgg gccatagccc tgatagacgg tttttcgccc 3420
tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact 3480
caaccctatc tcggtctatt cttttgattt ataagggatt ttgccgattt cggcctattg 3540
gttaaaaaat gagctgattt aacaaaaatt taacgcgaat tttaacaaaa tattaacgct 3600
tacaatttag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc 3660
taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa 3720
tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt 3780
gcggcatttt gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct 3840
gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc 3900
cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta 3960
tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac 4020
tattctcaga atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc 4080
atgacagtaa gagaattatg cagtgctgcc ataaccatga gtgataacac tgcggccaac 4140
ttacttctga caacgatcgg aggaccgaag gagctaaccg ctttttgca caacatgggg 4200
gatcatgtaa ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac 4260
gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc 4320
gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt 4380
gcaggaccac ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga 4440
gccggtgagc gtggttctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc 4500
cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag 4560
atcgctgaga taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca 4620
tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc 4680
ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca 4740
gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc 4800
tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta 4860
ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt 4920
ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc 4980
gctctgctaa tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg 5040
ttggactcaa gacgatagtt accggataag gcgcagcggt cgggctgaac gggggttcg 5100
tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag 5160
ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc 5220
agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat 5280
agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg 5340
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 5400
tggccttttg ctcacat 5417

<210> 273

<211> 11157
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS11979

<400> 273
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg ccaagtctgg catctaccag 2460
attaagaaca ccctgaataa caaagtgtat gttggatctg ccaaggactt tgagaaaaga 2520
tggaagaggc acttcaagga tctggaaaag ggttgccaca gctccatcaa gctccaaagg 2580
tctttcaaca aacatggcaa tgtctttgag tgctccattt tggaagagat cccttatgaa 2640
aaggacctca ttatagagag ggaaaacttc tggatcaagg agctgaactc aaaaatcaat 2700
ggatacaaca ttgcagatgc cacatttggt gacacctgtt ctactcaccc ccttaaggaa 2760
gagatcatta aaaagaggtc tgagactgtg aaggctaaga tgttgaaact ggggcctgat 2820
ggcagaaagg cactctactc taagccaggt agcaaaaatg gaaggtggaa ccctgagacc 2880
cacaagttct gcaaatgtgg ggtgagaata cagacatcag cctacacttg cagcaagtgt 2940
aggaacagat caggagagaa taactccttt ttcaaccaca agcattctga catcaccaag 3000
tccaaaattt caggcaagat gaaaggtaag aaaccatcta actccggagt ggatctacgc 3060
acgctcggct acagccagca gcaacaggag aagatcaaac cgaaggttcg ttcgacagtg 3120
gcgcagcacc acgaggcact ggtcggccac gggtttacac acgcgcacat cgttgcgtta 3180
agccaacacc cggcagcgtt agggaccgtc gctgtcaagt atcaggacat gatcgcagcg 3240
ttgccagagg cgacacacga agcgatcgtt ggcgtcggca aacagtggtc cggcgcacgc 3300
gctctggagg ccttgctcac ggtggcggga gagttgagag gtccaccgtt acagttggac 3360
acaggccaac ttctcaagat tgcaaaacgt ggcggcgtga ccgcagtgga ggcagtgcat 3420
gcatggcgca atgcactgac gggtgccccg ctcaacttga ccccggagca ggtggtggcc 3480
atcgccagcc acgatggcgg caagcaggcg ctggagacgg tccagcggct gttgccggtg 3540
ctgtgccagg cccacggctt gacccccag caggtggtgg ccatcgccag caatggcggt 3600
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc 3660
ttgaccccgg agcaggtggt ggccatcgcc agcaatattg gtggcaagca ggcgctggag 3720
acggtgcagg cgctgttgcc ggtgctgtgc caggcccacg gcttgacccc ccagcaggtg 3780

gtggccatcg ccagcaatgg cggtggcaag caggcgctgg agacggtcca gcggctgttg 3840
ccggtgctgt gccaggccca cggcttgacc ccggagcagg tggtggccat cgccagcaat 3900
attggtggca agcaggcgct ggagacggtg caggcgctgt tgccggtgct gtgccaggcc 3960
cacggcttga ccccggagca ggtggtggcc atcgccagca atattggtgg caagcaggcg 4020
ctggagacgg tgcaggcgct gttgccggtg ctgtgccagg cccacggctt gaccccggag 4080
caggtggtgg ccatcgccag caatattggt ggcaagcagg cgctggagac ggtgcaggcg 4140
ctgttgccgg tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc 4200
agccacgatg gcggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc 4260
caggcccacg gcttgacccc ggagcaggtg gtggccatcg ccagccacga tggcggcaag 4320
caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc 4380
ccccagcagg tggtggccat cgccagcaat ggcggtggca agcaggcgct ggagacggtc 4440
cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc 4500
atcgccagca atattggtgg caagcaggcg ctggagacgg tgcaggcgct gttgccggtg 4560
ctgtgccagg cccacggctt gaccccggag caggtggtgg ccatcgccag caatattggt 4620
ggcaagcagg cgctggagac ggtgcaggcg ctgttgccgg tgctgtgcca ggcccacggc 4680
ttgaccccgg agcaggtggt ggccatcgcc agccacgatg gcggcaagca ggcgctggag 4740
acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg 4800
gtggccatcg ccagccacga tggcggcaag caggcgctgg agacggtcca gcggctgttg 4860
ccggtgctgt gccaggccca cggcttgacc ccggagcagg tggtggccat cgccagccac 4920
gatggcggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc 4980
cacggcttga ccccccagca ggtggtggcc atcgccagca atggcggtgg caagcaggcg 5040
ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt gaccccggag 5100
caggtggtgg ccatcgccag ccacgatggc ggcaagcagg cgctggagac ggtccagcgg 5160
ctgttgccgg tgctgtgcca ggcccacggc ttgacccctc agcaggtggt ggccatcgcc 5220
agcaatggcg gcggcaggcc ggcgctggag agcattgttg cccagttatc tcgccctgat 5280
ccgtcggccg actgataact cgagcgatcc tctagacgag ctcctcgagc ctgcagcagc 5340
tgaagctttg gacttcttcg ccagaggttt ggtcaagtct ccaatcaagg ttgtcggctt 5400
gtctaccttg ccagaaattt acgaaaagat ggaaaagggt caaatcgttg gtagatacgt 5460
tgttgacact tctaaataag cgaatttctt atgatttatg atttttatta ttaaataagt 5520
tataaaaaaa ataagtgtat acaaatttta aagtgactct taggttttaa aacgaaaatt 5580
cttattcttg agtaactctt tcctgtaggt caggttgctt tctcaggtat agcatgaggt 5640
cgctcttatt gaccacacct ctaccggcat gcaagcttgg cgtaatcatg gtcatagctg 5700
tttcctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata 5760
aagtgtaaag cctggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca 5820
ctgcccgctt tccagtcggg aaacctgtcg tgccagcaga tctattacat tatgggtggt 5880
atgttggaat aaaaatcaac tatcatctac taactagtat ttacgttact agtatattat 5940
catatacggt gttagaagat gacgcaaatg atgagaaata gtcatctaaa ttagtggaag 6000
ctgaaacgca aggattgata atgtaatagg atcaatgaat attaacatat aaaatgatga 6060
taataatatt tatagaattg tgtagaattg cagattccct tttatggatt cctaaatcct 6120
cgaggagaac ttctagtata tctacatacc taatattatt gccttattaa aaatggaatc 6180
ccaacaatta catcaaaatc cacattctct tcaaaatcaa ttgtcctgta cttccttgtt 6240
catgtgtgtt caaaaacgtt atatttatag gataattata ctctatttct caacaagtaa 6300
ttggttgttt ggccgagcgg tctaaggcgc ctgattcaag aaatatcttg accgcagtta 6360
actgtgggaa tactcaggta tcgtaagatg caagagttcg aatctcttag caaccattat 6420
ttttttcctc aacataacga gaacacacag gggcgctatc gcacagaatc aaattcgatg 6480
actggaaatt ttttgttaat ttcagaggtc gcctgacgca tatacctttt tcaactgaaa 6540
aattgggaga aaaaggaaag gtgagagccg cggaaccggc ttttcatata gaatagagaa 6600
gcgttcatga ctaaatgctt gcatcacaat acttgaagtt gacaatatta tttaaggacc 6660
tattgttttt tccaataggt ggttagcaat cgtcttactt tctaactttt cttacctttt 6720
acatttcagc aatatatata tatatatttc aaggatatac cattctaatg tctgccccta 6780
agaagatcgt cgttttgcca ggtgaccacg ttggtcaaga aatcacagcc gaagccatta 6840
aggttcttaa agctatttct gatgttcgtt ccaatgtcaa gttcgatttc gaaaatcatt 6900
taattggtgg tgctgctatc gatgctacag gtgtcccact tccagatgag gcgctggaag 6960
cctccaagaa ggttgatgcc gttttgttag gtgctgtggg tggtcctaaa tggggtaccg 7020
gtagtgttag acctgaacaa ggtttactaa aaatccgtaa agaacttcaa ttgtacgcca 7080
acttaagacc atgtaacttt gcatccgact ctcttttaga cttatctcca atcaagccac 7140
aatttgctaa aggtactgac ttcgttgttg tcagagaatt agtgggaggt atttactttg 7200
gtaagagaaa ggaagacgat ggtgatggtg tcgcttggga tagtgaacaa tacaccgttc 7260
cagaagtgca aagaatcaca agaatggccg ctttcatggc cctacaacat gagccaccat 7320
tgcctatttg gtccttggat aaagctaatg ttttggcctc ttcaagatta tggagaaaaa 7380
ctgtggagga aaccatcaag aacgaattcc ctacattgaa ggttcaacat caattgattg 7440
attctgccgc catgatccta gttaagaacc caacccacct aaatggtatt ataatcacca 7500
gcaacatgtt tggtgatatc atctccgatg aagcctccgt tatcccaggt tccttgggtt 7560
tgttgccatc tgcgtccttg gcctctttgc cagacaagaa caccgcattt ggtttgtacg 7620
aaccatgcca cggttctgct ccagatttgc caaagaataa ggtcaaccct atcgccacta 7680
tcttgtctgc tgcaatgatg ttgaaattgt cattgaactt gcctgaagaa ggtaaggcca 7740
ttgaagatgc agttaaaaag gttttggatg caggtatcag aactggtgat ttaggtggtt 7800
ccaacagtac cacggaagtc ggtgatgctg tcgccgaaga agttaagaaa atccttgctt 7860

aaaaagattc tctttttta tgatatttgt acataaactt tataaatgaa attcataata 7920
gaaacgacac gaaattacaa aatggaatat gttcataggg tagacgaaac tatatacgca 7980
atctacatac atttatcaag aaggagaaaa aggaggatgt aaaggaatac aggtaagcaa 8040
attgatacta atggctcaac gtgataagga aaaagaattg cactttaaca ttaatattga 8100
caaggaggag ggcaccacac aaaaagttag gtgtaacaga aaatcatgaa actatgattc 8160
ctaatttata tattggagga ttttctctaa aaaaaaaaaa atacaacaaa taaaaaacac 8220
tcaatgacct gaccatttga tggagtttaa gtcaatacct tcttgaacca tttcccataa 8280
tggtgaaagt tccctcaaga attttactct gtcagaaacg gccttaacga cgtagtcgac 8340
ctcctcttca gtactaaatc taccaatacc aaatctgatg gaagaatggg ctaatgcatc 8400
atccttaccc agcgcatgta aaacataaga aggttctagg gaagcagatg tacaggctga 8460
acccgaggat aatgcgatat cccttagtgc catcaataaa gattctcctt ccacgtaggc 8520
gaaagaaacg ttaacacacc ctggataacg atgatctgga gatccgttca acgtggtatg 8580
ttcagcggat aatagacctt tgactaattt atcggatagt cttttgatgt gagcttggtc 8640
gttgtcaaat tctttcttca tcaatctcgc agcttcacca aatcccgcta ccaatggggg 8700
ggccaaagta ccagatctgc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt 8760
gcgtattggg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct 8820
gcggcgagcg gtatcagcat cgatgaattc cacggactat agactatact agtatactcc 8880
gtctactgta cgatacactt ccgctcaggt ccttgtcctt taacgaggcc ttaccactct 8940
tttgttactc tattgatcca gctcagcaaa ggcagtgtga tctaagattc tatcttcgcg 9000
atgtagtaaa actagctaga ccgagaaaga gactagaaat gcaaaaggca cttctacaat 9060
ggctgccatc attattatcc gatgtgacgc tgcagcttct caatgatatt cgaatacgct 9120
ttgaggagat acagcctaat atccgacaaa ctgttttaca gatttacgat cgtacttgtt 9180
acccatcatt gaattttgaa catccgaacc tgggagtttt ccctgaaaca gatagtatat 9240
ttgaacctgt ataataatat atagtctagc gctttacgga agacaatgta tgtatttcgg 9300
ttcctggaga aactattgca tctattgcat aggtaatctt gcacgtcgca tccccggttc 9360
attttctgcg tttccatctt gcacttcaat agcatatctt tgttaacgaa gcatctgtgc 9420
ttcattttgt agaacaaaaa tgcaacgcga gagcgctaat ttttcaaaca aagaatctga 9480
gctgcatttt tacagaacag aaatgcaacg cgaaagcgct attttaccaa cgaagaatct 9540
gtgcttcatt tttgtaaaac aaaaatgcaa cgcgagagcg ctaatttttc aaacaaagaa 9600
tctgagctgc atttttacag aacagaaatg caacgcgaga gcgctatttt accaacaaag 9660
aatctatact tctttttgt tctacaaaaa tgcatcccga gagcgctatt tttctaacaa 9720
agcatcttag attacttttt ttctcctttg tgcgctctat aatgcagtct cttgataact 9780
ttttgcactg taggtccgtt aaggttagaa gaaggctact ttggtgtcta ttttctcttc 9840
cataaaaaaa gcctgactcc acttcccgcg tttactgatt actagcgaag ctgcgggtgc 9900
atttttttcaa gataaaggca tccccgatta tattctatac cgatgtggat tgcgcatact 9960
ttgtgaacag aaagtgatag cgttgatgat tcttcattgg tcagaaaatt atgaacggtt 10020
tcttctattt tgtctctata tactacgtat aggaaatgtt tacattttcg tattgttttc 10080
gattcactct atgaatagtt cttactacaa tttttttgtc taaagagtaa tactagagat 10140
aaacataaaa aatgtagagg tcgagtttag atgcaagttc aaggagcgaa aggtggatgg 10200
gtaggttata tagggatata gcacagagat atatagcaaa gagatacttt tgagcaatgt 10260
ttgtggaagc ggtattcgca atattttagt agctcgttac agtccggtgc gtttttggtt 10320
ttttgaaagt gcgtcttcag agcgcttttg gttttcaaaa gcgctctgaa gttcctatac 10380
tttctagaga ataggaactt cggaatagga acttcaaagc gtttccgaaa acgagcgctt 10440
ccgaaaatgc aacgcgagct gcgcacatac agctcactgt tcacgtcgca cctatatctg 10500
cgtgttgcct gtatatatat atacatgaga agaacggcat agtgcgtgtt tatgcttaaa 10560
tgcgtactta tatgcgtcta tttatgtagg atgaaaggta gtctagtacc tcctgtgata 10620
ttatcccatt ccatgcgggg tatcgtatgc ttccttcagc actacccttt agctgttcta 10680
tatgctgcca ctcctcaatt ggattagtct catccttcaa tgctatcatt tcctttgata 10740
ttggatcata tgcatagtac cgagaaacta gtgcgaagta gtgatcaggt attgctgtta 10800
tctgatgagt atacgttgtc ctggccacgg cagaagcacg cttatcgctc caatttccca 10860
caacattagt caactccgtt aggcccttca ttgaaagaaa tgaggtcatc aaatgtcttc 10920
caatgtgaga ttttgggcca ttttttatag caaagattga ataaggcgca tttttcttca 10980
aagctttatt gtacgatctg actaagttat cttttaataa ttggtattcc tgtttattgc 11040
ttgaagaatt gccggtccta tttactcgtt ttaggactgg ttcagaattc atcgatgctc 11100
actcaaaggt cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaac 11157

<210> 274

<211> 11151
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS8521

<400> 274
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60

ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa atagggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg ccaagtctgg catctaccag 2460
attaagaaca ccctgaataa caaagtgtat gttggatctg ccaaggactt tgagaaaaga 2520
tggaagaggc acttcaagga tctggaaaag ggttgccaca gctccatcaa gctccaaagg 2580
tctttcaaca aacatggcaa tgtctttgag tgctccattt tggaagagat cccttatgaa 2640
aaggacctca ttatagagag ggaaaacttc tggatcaagg agctgaactc aaaaatcaat 2700
ggatacaaca ttgcagatgc cacatttggt gacacctgtt ctactcaccc ccttaaggaa 2760
gagatcatta aaaagaggtc tgagactgtg aaggctaaga tgttgaaact ggggcctgat 2820
ggcagaaagg cactctactc taagccaggt agcaaaaatg gaaggtggaa ccctgagacc 2880
cacaagttct gcaaatgtgg ggtgagaata cagacatcag cctacacttg cagcaagtgt 2940
aggaacagat caggagagaa taactccttt ttcaaccaca agcattctga catcaccaag 3000
tccaaaattt caggcaagat gaaaggtcaa ggtccatccg gagtggatct acgcacgctc 3060
ggctacagcc agcagcaaca ggagaagatc aaaccgaagg ttcgttcgac agtggcgcag 3120
caccacgagg cactggtcgg ccacgggttt acacacgcgc acatcgttgc gttaagccaa 3180
cacccggcag cgttagggac cgtcgctgtc aagtatcagg acatgatcgc agcgttgcca 3240
gaggcgacac acgaagcgat cgttggcgtc ggcaaacagt ggtccggcgc acgcgctctg 3300
gaggccttgc tcacggtggc gggagagttg agaggtccac cgttacagtt ggacacaggc 3360
caacttctca agattgcaaa acgtggcggc gtgaccgcag tggaggcagt gcatgcatgg 3420
cgcaatgcac tgacgggtgc cccgctcaac ttgaccccgg agcaggtggt ggccatcgcc 3480
agccacgatg gcggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc 3540
caggcccacg gcttgacccc ccagcaggtg gtggccatcg ccagcaatgg cggtggcaag 3600
caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc 3660
ccggagcagg tggtggccat cgccagcaat attggtggca agcaggcgct ggagacggtg 3720
caggcgctgt tgccggtgct gtgccaggcc cacggcttga ccccccagca ggtggtggcc 3780
atcgccagca atggcggtgg caagcaggcg ctggagacgg tccagcggct gttgccggtg 3840
ctgtgccagg cccacggctt gaccccggag caggtggtgg ccatcgccag caatattggt 3900
ggcaagcagg cgctggagac ggtgcaggcg ctgttgccgg tgctgtgcca ggcccacggc 3960
ttgaccccgg agcaggtggt ggccatcgcc agcaatattg gtggcaagca ggcgctggag 4020
acggtgcagg cgctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg 4080
gtggccatcg ccagcaatat tggtggcaag caggcgctgg agacggtgca ggcgctgttg 4140

```
ccggtgctgt gccaggccca cggcttgacc ccggagcagg tggtggccat cgccagccac   4200
gatggcggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc   4260
cacggcttga ccccggagca ggtggtggcc atcgccagcc acgatggcgg caagcaggcg   4320
ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt gacccccag    4380
caggtggtgg ccatcgccag caatggcggt ggcaagcagg cgctggagac ggtccagcgg   4440
ctgttgccgg tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc   4500
agcaatattg gtggcaagca ggcgctggag acggtgcagg cgctgttgcc ggtgctgtgc   4560
caggcccacg gcttgacccc ggagcaggtg gtggccatcg ccagcaatat tggtggcaag   4620
caggcgctgg agacggtgca ggcgctgttg ccggtgctgt gccaggccca cggcttgacc   4680
ccggagcagg tggtggccat cgccagccac gatggcggca agcaggcgct ggagacggtc   4740
cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc   4800
atcgccagcc acgatggcgg caagcaggcg ctggagacgg tccagcggct gttgccggtg   4860
ctgtgccagg cccacggctt gaccccggag caggtggtgg ccatcgccag ccacgatggc   4920
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc   4980
ttgacccccc agcaggtggt ggccatcgcc agcaatggcg gtggcaagca ggcgctggag   5040
acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg   5100
gtggccatcg ccagccacga tggcggcaag caggcgctgg agacggtcca gcggctgttg   5160
ccggtgctgt gccaggccca cggcttgacc cctcagcagg tggtggccat cgccagcaat   5220
ggcggcggca ggccggcgct ggagagcatt gttgcccagt tatctcgccc tgatccgtcg   5280
gccgactgat aactcgagcg atcctctaga cgagctcctc gagcctgcag cagctgaagc   5340
tttggacttc ttcgccagag gtttggtcaa gtctccaatc aaggttgtcg gcttgtctac   5400
cttgccagaa atttacgaaa agatggaaaa gggtcaaatc gttggtagat acgttgttga   5460
cacttctaaa taagcgaatt tcttatgatt tatgattttt attattaaat aagttataaa   5520
aaaaataagt gtatacaaat tttaaagtga ctcttaggtt ttaaaacgaa aattcttatt   5580
cttgagtaac tctttcctgt aggtcaggtt gctttctcag gtatagcatg aggtcgctct   5640
tattgaccac acctctaccg gcatgcaagc ttggcgtaat catggtcata gctgtttcct   5700
gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag cataaagtgt   5760
aaagcctggg gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc   5820
gctttccagt cgggaaacct gtcgtgccag cagatctatt acattatggg tggtatgttg   5880
gaataaaaat caactatcat ctactaacta gtatttacgt tactagtata ttatcatata   5940
cggtgttaga agatgacgca aatgatgaga aatagtcatc taaattagtg gaagctgaaa   6000
cgcaaggatt gataatgtaa taggatcaat gaatattaac atataaaatg atgataataa   6060
tatttataga attgtgtaga attgcagatt ccctttatg gattcctaaa tcctcgagga   6120
gaacttctag tatatctaca tacctaatat tattgcctta ttaaaaatgg aatcccaaca   6180
attacatcaa aatccacatt ctcttcaaaa tcaattgtcc tgtacttcct tgttcatgtg   6240
tgttcaaaaa cgttatattt ataggataat tatactctat ttctcaacaa gtaattggtt   6300
gtttggccga gcggtctaag gcgcctgatt caagaaatat cttgaccgca gttaactgtg   6360
ggaatactca ggtatcgtaa gatgcaagag ttcgaatctc ttagcaacca ttatttttt    6420
cctcaacata acgagaacac acaggggcgc tatcgcacag aatcaaattc gatgactgga   6480
aatttttgt taatttcaga ggtcgcctga cgcatatacc tttttcaact gaaaaattgg    6540
gagaaaaagg aaaggtgaga gccgcggaac cggcttttca tatagaatag agaagcgttc   6600
atgactaaat gcttgcatca caatacttga agttgacaat attatttaag gacctattgt   6660
tttttccaat aggtggttag caatcgtctt actttctaac ttttcttacc ttttacattt   6720
cagcaatata tatatatata tttcaaggat ataccattct aatgtctgcc cctaagaaga   6780
tcgtcgtttt gccaggtgac cacgttggtc aagaaatcac agccgaagcc attaaggttc   6840
ttaaagctat ttctgatgtt cgttccaatg tcaagttcga tttcgaaaat catttaattg   6900
gtggtgctgc tatcgatgct acaggtgtcc cacttccaga tgaggcgctg gaagcctcca   6960
agaaggttga tgccgttttg ttaggtgctg tgggtggtcc taaatggggt accggtagtg   7020
ttagacctga acaaggttta ctaaaaatcc gtaaagaact tcaattgtac gccaacttaa   7080
gaccatgtaa ctttgcatcc gactctcttt tagacttatc tccaatcaag ccacaatttg   7140
ctaaaggtac tgacttcgtt gttgtcagag aattagtggg aggtatttac tttggtaaga   7200
gaaaggaaga cgatggtgat ggtgtcgctt gggatagtga acaatacacc gttccagaag   7260
tgcaaagaat cacaagaatg gccgctttca tggccctaca acatgagcca ccattgccta   7320
tttggtcctt ggataaagct aatgttttgg cctcttcaag attatggaga aaaactgtgg   7380
aggaaaccat caagaacgaa ttccctacat tgaaggttca acatcaattg attgattctg   7440
ccgccatgat cctagttaag aacccaaccc acctaaatgg tattataatc accagcaaca   7500
tgtttggtga tatcatctcc gatgaagcct ccgttatccc aggttccttg ggtttgttgc   7560
catctgcgtc cttggcctct ttgccagaca agaacaccgc atttggtttg tacgaaccat   7620
gccacggttc tgctccagat ttgccaaaga ataaggtcaa ccctatcgcc actatcttgt   7680
ctgctgcaat gatgttgaaa ttgtcattga acttgcctga agaaggtaag gccattgaag   7740
atgcagttaa aaaggttttg gatgcaggta tcagaactgg tgatttaggt ggttccaaca   7800
gtaccacgga agtcggtgat gctgtcgccg aagaagttaa gaaaatcctt gcttaaaaag   7860
attctctttt tttatgatat ttgtacataa actttataaa tgaaattcat aatagaaacg   7920
acacgaaatt acaaaatgga atatgttcat agggtagacg aaactatata cgcaatctac   7980
atacatttat caagaaggag aaaaaggagg atgtaaagga atacaggtaa gcaaattgat   8040
actaatggct caacgtgata aggaaaaaga attgcacttt aacattaata ttgacaagga   8100
ggagggcacc acacaaaaag ttaggtgtaa cagaaaatca tgaaactatg attcctaatt   8160
tatatattgg aggattttct ctaaaaaaaa aaaaatacaa caaataaaaa acactcaatg   8220
```

acctgaccat ttgatggagt ttaagtcaat accttcttga accatttccc ataatggtga 8280
aagttccctc aagaatttta ctctgtcaga aacggcctta acgacgtagt cgacctcctc 8340
ttcagtacta aatctaccaa taccaaatct gatggaagaa tgggctaatg catcatcctt 8400
acccagcgca tgtaaaacat aagaaggttc tagggaagca gatgtacagg ctgaacccga 8460
ggataatgcg atatccctta gtgccatcaa taaagattct ccttccacgt aggcgaaaga 8520
aacgttaaca caccctggat aacgatgatc tggagatccg ttcaacgtgg tatgttcagc 8580
ggataataga cctttgacta atttatcgga tagtcttttg atgtgagctt ggtcgttgtc 8640
aaattctttc ttcatcaatc tcgcagcttc accaaatccc gctaccaatg gggggccaa 8700
agtaccagat ctgctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat 8760
tgggcgctct tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc ggctgcggcg 8820
agcggtatca gcatcgatga attccacgga ctatagacta tactagtata ctccgtctac 8880
tgtacgatac acttccgctc aggtccttgt cctttaacga ggccttacca ctcttttgtt 8940
actctattga tccagctcag caaaggcagt gtgatctaag attctatctt cgcgatgtag 9000
taaaactagc tagaccgaga aagagactag aaatgcaaaa ggcacttcta caatggctgc 9060
catcattatt atccgatgtg acgctgcagc ttctcaatga tattcgaata cgctttgagg 9120
agatacagcc taatatccga caaactgttt tacagattta cgatcgtact tgttacccat 9180
cattgaattt tgaacatccg aacctgggag ttttccctga aacagatagt atatttgaac 9240
ctgtataata atatatagtc tagcgcttta cggaagacaa tgtatgtatt tcggttcctg 9300
gagaaactat tgcatctatt gcataggtaa tcttgcacgt cgcatccccg gttcattttc 9360
tgcgtttcca tcttgcactt caatagcata tctttgttaa cgaagcatct gtgcttcatt 9420
ttgtagaaca aaaatgcaac gcgagagcgc taatttttca aacaaagaat ctgagctgca 9480
tttttacaga acagaaatgc aacgcgaaag cgctatttta ccaacgaaga atctgtgctt 9540
cattttgta aaacaaaaat gcaacgcgag agcgctaatt tttcaaacaa agaatctgag 9600
ctgcattttt acagaacaga aatgcaacgc gagagcgcta ttttaccaac aaagaatcta 9660
tacttctttt ttgttctaca aaaatgcatc ccgagagcgc tatttttcta acaaagcatc 9720
ttagattact ttttttctcc tttgtgcgct ctataatgca gtctcttgat aacttttgc 9780
actgtaggtc cgttaaggtt agaagaaggc tactttggtg tctattttct cttccataaa 9840
aaaagcctga ctccacttcc cgcgtttact gattactagc gaagctgcgg gtgcattttt 9900
tcaagataaa ggcatccccg attatattct ataccgatgt ggattgcgca tactttgtga 9960
acagaaagtg atagcgttga tgattcttca ttggtcagaa aattatgaac ggtttcttct 10020
attttgtctc tatatactac gtataggaaa tgtttacatt ttcgtattgt tttcgattca 10080
ctctatgaat agttcttact acaatttttt tgtctaaaga gtaatactag agataaacat 10140
aaaaaatgta gaggtcgagt ttagatgcaa gttcaaggag cgaaaggtgg atgggtaggt 10200
tatataggga tatagcacag agatatatag caaagagata cttttgagca atgtttgtgg 10260
aagcggtatt cgcaatattt tagtagctcg ttacagtccg gtgcgttttt ggtttttga 10320
aagtgcgtct tcagagcgct tttggttttc aaaagcgctc tgaagttcct atactttcta 10380
gagaatagga acttcggaat aggaacttca aagcgtttcc gaaaacgagc gcttccgaaa 10440
atgcaacgcg agctgcgcac atacagctca ctgttcacgt cgcacctata tctgcgtgtt 10500
gcctgtatat atatatacat gagaagaacg gcatagtgcg tgtttatgct taaatgcgta 10560
cttatatgcg tctatttatg taggatgaaa ggtagtctag tacctcctgt gatattatcc 10620
cattccatgc ggggtatcgt atgcttcctt cagcactacc ctttagctgt tctatatgct 10680
gccactcctc aattggatta gtctcatcct tcaatgctat catttccttt gatattggat 10740
catatgcata gtaccgagaa actagtgcga agtagtgatc aggtattgct gttatctgat 10800
gagtatacgt tgtcctggcc acggcagaag cacgcttatc gctccaattt cccacaacat 10860
tagtcaactc cgttaggccc ttcattgaaa gaaatgaggt catcaaatgt cttccaatgt 10920
gagattttgg gccatttttt atagcaaaga ttgaataagg cgcatttttc ttcaaagctt 10980
tattgtacga tctgactaag ttatctttta ataattggta ttcctgttta ttgcttgaag 11040
aattgccggt cctatttact cgttttagga ctggttcaga attcatcgat gctcactcaa 11100
aggtcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa c 11151

<210> 275

<211> 10728
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS12100

<400> 275
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420

aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg ccaagtctgg catctaccag 2460
attaagaaca ccctgaataa caaagtgtat gttggatctg ccaaggactt tgagaaaaga 2520
tggaagaggc acttcaagga tctggaaaag ggttgccaca gctccatcaa gctccaaagg 2580
tctttcaaca aacatggcaa tgtctttgag tgctccattt tggaagagat cccttatgaa 2640
aaggacctca ttatagagag ggaaaacttc tggatcaagg agctgaactc aaaaatcaat 2700
ggatacaaca ttgcagatgc cacatttggt gacacctgtt ctactcaccc ccttaaggaa 2760
gagatcatta aaaagaggtc tgagactgtg aaggctaaga tgttgaaaca aggtccatcc 2820
ggagtggatc tacgcacgct cggctacagc cagcagcaac aggagaagat caaaccgaag 2880
gttcgttcga cagtggcgca gcaccacgag gcactggtcg gccacgggtt tacacacgcg 2940
cacatcgttg cgttaagcca acacccggca gcgttaggga ccgtcgctgt caagtatcag 3000
gacatgatcg cagcgttgcc agaggcgaca cacgaagcga tcgttggcgt cggcaaacag 3060
tggtccggcg cacgcgctct ggaggccttg ctcacggtgg cgggagagtt gagaggtcca 3120
ccgttacagt tggacacagg ccaacttctc aagattgcaa aacgtggcgg cgtgaccgca 3180
gtggaggcag tgcatgcatg gcgcaatgca ctgacgggtg ccccgctcaa cttgaccccc 3240
cagcaggtgg tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag 3300
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 3360
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3420
tgccaggccc acggcttgac cccccagcag gtggtggcca tcgccagcaa tggcggtggc 3480
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 3540
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 3600
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 3660
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 3720
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 3780
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 3840
ggcttgaccc cccagcaggt ggtggccatc gccagcaata atggtggcaa gcaggcgctg 3900
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 3960
gtggtggcca tcgccagcaa taatggtggc aagcaggcgc tggagacggt ccagcggctg 4020
ttgccggtgc tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc 4080
aatggcggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 4140
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 4200
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 4260
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 4320
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 4380
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 4440
tgccaggccc acggcttgac cccccagcag gtggtggcca tcgccagcaa tggcggtggc 4500

```
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg   4560
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg   4620
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg   4680
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg   4740
gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg tggccatcgc cagcaatggc   4800
ggcggcaggc cggcgctgga gagcattgtt gcccagttat ctcgccctga tccgtcggcc   4860
gactgataac tcgagcgatc ctctagacga gctcctcgag cctgcagcag ctgaagcttt   4920
ggacttcttc gccagaggtt tggtcaagtc tccaatcaag gttgtcggct tgtctacctt   4980
gccagaaatt tacgaaaaga tggaaaaggg tcaaatcgtt ggtagatacg ttgttgacac   5040
ttctaaataa gcgaatttct tatgatttat gatttttatt attaaataag ttataaaaaa   5100
aataagtgta tacaaatttt aaagtgactc ttaggtttta aaacgaaaat tcttattctt   5160
gagtaactct ttcctgtagg tcaggttgct ttctcaggta tagcatgagg tcgctcttat   5220
tgaccacacc tctaccggca tgcaagcttg gcgtaatcat ggtcatagct gtttcctgtg   5280
tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat aaagtgtaaa   5340
gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgctc actgcccgct   5400
ttccagtcgg gaaacctgtc gtgccagcag atctattaca ttatgggtgg tatgttggaa   5460
taaaaatcaa ctatcatcta ctaactagta tttacgttac tagtatatta tcatatacgg   5520
tgttagaaga tgacgcaaat gatgagaaat agtcatctaa attagtggaa gctgaaacgc   5580
aaggattgat aatgtaatag gatcaatgaa tattaacata taaaatgatg ataataatat   5640
ttatagaatt gtgtagaatt gcagattccc ttttatggat tcctaaatcc tcgaggagaa   5700
cttctagtat atctacatac ctaatattat tgccttatta aaaatggaat cccaacaatt   5760
acatcaaaat ccacattctc ttcaaaatca attgtcctgt acttccttgt tcatgtgtgt   5820
tcaaaaacgt tatatttata ggataattat actctatttc tcaacaagta attggttgtt   5880
tggccgagcg gtctaaggcg cctgattcaa gaaatatctt gaccgcagtt aactgtggga   5940
atactcaggt atcgtaagat gcaagagttc gaatctctta gcaaccatta ttttttcct   6000
caacataacg agaacacaca ggggcgctat cgcacagaat caaattcgat gactggaaat   6060
tttttgttaa tttcagaggt cgcctgacgc atatacctttt ttcaactgaa aaattgggag   6120
aaaaaggaaa ggtgagagcc gcggaaccgg cttttcatat agaatagaga agcgttcatg   6180
actaaatgct tgcatcacaa tacttgaagt tgacaatatt atttaaggac ctattgtttt   6240
ttccaatagg tggttagcaa tcgtcttact ttctaacttt tcttaccttt tacatttcag   6300
caatatatat atatatattt caaggatata ccattctaat gtctgcccct aagaagatcg   6360
tcgttttgcc aggtgaccac gttggtcaag aaatcacagc cgaagccatt aaggttctta   6420
aagctatttc tgatgttcgt tccaatgtca agttcgattt cgaaaatcat ttaattggtg   6480
gtgctgctat cgatgctaca ggtgtcccac ttccagatga ggcgctggaa gcctccaaga   6540
aggttgatgc cgttttgtta ggtgctgtgg gtggtcctaa atggggtacc ggtagtgtta   6600
gacctgaaca aggtttacta aaaatccgta aagaacttca attgtacgcc aacttaagac   6660
catgtaactt tgcatccgac tctcttttag acttatctcc aatcaagcca caatttgcta   6720
aaggtactga cttcgttgtt gtcagagaat tagtgggagg tatttacttt ggtaagagaa   6780
aggaagacga tggtgatggt gtcgcttggg atagtgaaca atacaccgtt ccagaagtgc   6840
aaagaatcac aagaatggcc gctttcatgg ccctacaaca tgagccacca ttgcctattt   6900
ggtccttgga taaagctaat gttttggcct cttcaagatt atggagaaaa actgtggagg   6960
aaaccatcaa gaacgaattc cctacattga aggttcaaca tcaattgatt gattctgccg   7020
ccatgatcct agttaagaac ccaacccacc taaatggtat tataatcacc agcaacatgt   7080
ttggtgatat catctccgat gaagcctccg ttatcccagg ttccttgggt ttgttgccat   7140
ctgcgtcctt ggcctctttg ccagacaaga acaccgcatt tggtttgtac gaaccatgcc   7200
acggttctgc tccagatttg ccaaagaata aggtcaaccc tatcgccact atcttgtctg   7260
ctgcaatgat gttgaaattg tcattgaact tgcctgaaga aggtaaggcc attgaagatg   7320
cagttaaaaa ggttttggat gcaggtatca gaactggtga tttaggtggt tccaacagta   7380
ccacggaagt cggtgatgct gtcgccgaag aagttaagaa aatccttgct taaaaagatt   7440
ctctttttttt atgatatttg tacataaact ttataaatga aattcataat agaaacgaca   7500
cgaaattaca aaatggaata tgttcatagg gtagacgaaa ctatatacgc aatctacata   7560
catttatcaa gaaggagaaa aaggaggatg taaaggaata caggtaagca aattgatact   7620
aatggctcaa cgtgataagg aaaaagaatt gcactttaac attaatattg acaaggagga   7680
gggcaccaca caaaaagtta ggtgtaacag aaaatcatga aactatgatt cctaatttat   7740
atattggagg attttctcta aaaaaaaaaa aatacaacaa ataaaaaaca ctcaatgacc   7800
tgaccatttg atggagttta agtcaatacc ttcttgaacc atttcccata atggtgaaag   7860
ttccctcaag aattttactc tgtcagaaac ggccttaacg acgtagtcga cctcctcttc   7920
agtactaaat ctaccaatac caaatctgat ggaagaatgg gctaatgcat catccttacc   7980
cagcgcatgt aaaacataag aaggttctag ggaagcagat gtacaggctg aacccgagga   8040
taatgcgata tcccttagtg ccatcaataa agattctcct tccacgtagg cgaaagaaac   8100
gttaacacac cctggataac gatgatctgg agatccgttc aacgtggtat gttcagcgga   8160
taatagacct ttgactaatt tatcggatag tcttttgatg tgagcttggt cgttgtcaaa   8220
ttctttcttc atcaatctcg cagcttcacc aaatcccgct accaatgggg gggccaaagt   8280
accagatctg ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg   8340
gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc   8400
ggtatcagca tcgatgaatt ccacggacta tagactatac tagtatactc cgtctactgt   8460
acgatacact tccgctcagg tccttgtcct ttaacgaggc cttaccactc ttttgttact   8520
ctattgatcc agctcagcaa aggcagtgtg atctaagatt ctatcttcgc gatgtagtaa   8580
```

aactagctag accgagaaag agactagaaa tgcaaaaggc acttctacaa tggctgccat 8640
cattattatc cgatgtgacg ctgcagcttc tcaatgatat tcgaatacgc tttgaggaga 8700
tacagcctaa tatccgacaa actgttttac agatttacga tcgtacttgt tacccatcat 8760
tgaattttga acatccgaac ctgggagttt tccctgaaac agatagtata tttgaacctg 8820
tataataata tatagtctag cgctttacgg aagacaatgt atgtatttcg gttcctggag 8880
aaactattgc atctattgca taggtaatct tgcacgtcgc atccccggtt cattttctgc 8940
gtttccatct tgcacttcaa tagcatatct ttgttaacga agcatctgtg cttcattttg 9000
tagaacaaaa atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt 9060
ttacagaaca gaaatgcaac gcgaaagcgc tattttacca acgaagaatc tgtgcttcat 9120
ttttgtaaaa caaaaatgca acgcgagagc gctaattttt caaacaaaga atctgagctg 9180
catttttaca gaacagaaat gcaacgcgag agcgctattt taccaacaaa gaatctatac 9240
ttctttttttg ttctacaaaa atgcatcccg agagcgctat ttttctaaca aagcatctta 9300
gattactttt tttctccttt gtgcgctcta taatgcagtc tcttgataac tttttgcact 9360
gtaggtccgt taaggttaga agaaggctac tttggtgtct attttctctt ccataaaaaa 9420
agcctgactc cacttcccgc gtttactgat tactagcgaa gctgcgggtg catttttttca 9480
agataaaggc atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca 9540
gaaagtgata gcgttgatga ttcttcattg gtcagaaaat tatgaacggt ttcttctatt 9600
ttgtctctat atactacgta taggaaatgt ttacattttc gtattgtttt cgattcactc 9660
tatgaatagt tcttactaca attttttttgt ctaaagagta atactagaga taaacataaa 9720
aaatgtagag gtcgagttta gatgcaagtt caaggagcga aaggtggatg ggtaggttat 9780
atagggatat agcacagaga tatatagcaa agagatactt ttgagcaatg tttgtggaag 9840
cggtattcgc aatattttag tagctcgtta cagtccggtg cgtttttggt tttttgaaag 9900
tgcgtcttca gagcgctttt ggttttcaaa agcgctctga agttcctata ctttctagag 9960
aataggaact tcggaatagg aacttcaaag cgtttccgaa aacgagcgct tccgaaaatg 10020
caacgcgagc tgcgcacata cagctcactg ttcacgtcgc acctatatct gcgtgttgcc 10080
tgtatatata tatacatgag aagaacggca tagtgcgtgt ttatgcttaa atgcgtactt 10140
atatgcgtct atttatgtag gatgaaaggt agtctagtac ctcctgtgat attatcccat 10200
tccatgcggg gtatcgtatg cttccttcag cactaccctt tagctgttct atatgctgcc 10260
actcctcaat tggattagtc tcatccttca atgctatcat ttcctttgat attggatcat 10320
atgcatagta ccgagaaact agtgcgaagt agtgatcagg tattgctgtt atctgatgag 10380
tatacgttgt cctggccacg gcagaagcac gcttatcgct ccaatttccc acaacattag 10440
tcaactccgt taggcccttc attgaaagaa atgaggtcat caaatgtctt ccaatgtgag 10500
attttgggcc atttttttata gcaaagattg aataaggcgc atttttcttc aaagctttat 10560
tgtacgatct gactaagtta tcttttaata attggtattc ctgtttattg cttgaagaat 10620
tgccggtcct atttactcgt tttaggactg gttcagaatt catcgatgct cactcaaagg 10680
tcggtaatac ggttatccac agaatcaggg gataacgcag gaaagaac 10728

<210> 276

<211> 1614
<212> DNA
<213> Artificial sequence

<220>
<223> NptIIT5-L RVD DNA insert

<400> 276
gcagcctcaa cttgaccccg gagcaggtgg tggccatcgc cagccacgat ggcggcaagc 60
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 120
cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg gagacggtcc 180
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag gtggtggcca 240
tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 300
tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc aatggcggtg 360
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 420
tgacccccca gcaggtggtg gccatcgcca gcaataatgg tggcaagcag gcgctggaga 480
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg 540
tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag cggctgttgc 600
cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc gccagcaata 660
atggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg tgccaggccc 720
acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc aagcaggcgc 780
tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg accccggagc 840
aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg gtgcaggcgc 900
tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg gccatcgcca 960
gcaataatgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc 1020
aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat ggcggcaagc 1080
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 1140
cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc 1200

agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag gtggtggcca 1260
tcgccagcaa taatggtggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 1320
tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc aatggcggtg 1380
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 1440
tgacccccca gcaggtggtg gccatcgcca gcaataatgg tggcaagcag gcgctggaga 1500
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg 1560
tggccatcgc cagcaatggc ggcggcaggc cggcgctgga gagcattgtg atgc 1614

<210> 277

<211> 1614
<212> DNA
<213> Artificial sequence

<220>
<223> NptIIT5-R RVD DNA insert

<400> 277
gcagcctcaa cttgaccccg gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc 60
aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc 120
cccagcaggt ggtggccatc gccagcaata atggtggcaa gcaggcgctg gagacggtcc 180
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag gtggtggcca 240
tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 300
tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc aatattggtg 360
gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag gcccacggct 420
tgacccccca gcaggtggtg gccatcgcca gcaataatgg tggcaagcag gcgctggaga 480
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg 540
tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag cggctgttgc 600
cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc gccagccacg 660
atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg tgccaggccc 720
acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc aagcaggcgc 780
tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg acccccccagc 840
aggtggtggc catcgccagc aataatggtg gcaagcaggc gctggagacg gtccagcggc 900
tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg gccatcgcca 960
gcaatggcgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc 1020
aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat ggcggcaagc 1080
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 1140
cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg gagacggtcc 1200
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag gtggtggcca 1260
tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 1320
tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc aatggcggtg 1380
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 1440
tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag gcgctggaga 1500
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg 1560
tggccatcgc cagcaatggc ggcggcaggc cggcgctgga gagcattgtg atgc 1614

<210> 278

<211> 1614
<212> DNA
<213> Artificial sequence

<220>
<223> NptIIT6-L RVD DNA insert

<400> 278
gcagcctcaa cttgaccccc cagcaggtgg tggccatcgc cagcaataat ggtggcaagc 60
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 120
cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg gagacggtgc 180
aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag gtggtggcca 240
tcgccagcaa tattggtggc aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc 300
tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc aataatggtg 360
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 420
tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag gcgctggaga 480
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg 540
tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag cggctgttgc 600

cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc gccagcaata 660
atggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg tgccaggccc 720
acggcttgac cccccagcag gtggtggcca tcgccagcaa taatggtggc aagcaggcgc 780
tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg accccggagc 840
aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg gtgcaggcgc 900
tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg gccatcgcca 960
gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg gtgctgtgcc 1020
aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaataat ggtggcaagc 1080
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 1140
cccagcaggt ggtggccatc gccagcaata atggtggcaa gcaggcgctg gagacggtcc 1200
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag gtggtggcca 1260
tcgccagcaa taatggtggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 1320
tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc aatattggtg 1380
gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag gcccacggct 1440
tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag gcgctggaga 1500
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg 1560
tggccatcgc cagcaatggc ggcggcaggc cggcgctgga gagcattgtg atgc 1614

<210> 279

<211> 1614
<212> DNA
<213> Artificial sequence

<220>
<223> NptIIT6-R RVD DNA insert

<400> 279
gcagcctcaa cttgaccccc cagcaggtgg tggccatcgc cagcaataat ggtggcaagc 60
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 120
cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg gagacggtgc 180
aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag gtggtggcca 240
tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 300
tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc aatattggtg 360
gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag gcccacggct 420
tgaccccca gcaggtggtg gccatcgcca gcaataatgg tggcaagcag gcgctggaga 480
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg 540
tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag cggctgttgc 600
cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc gccagcaata 660
ttggtggcaa gcaggcgctg gagacggtgc aggcgctgtt gccggtgctg tgccaggccc 720
acggcttgac cccccagcag gtggtggcca tcgccagcaa taatggtggc aagcaggcgc 780
tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg accccggagc 840
aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg gtgcaggcgc 900
tgttgccggt gctgtgccag gcccacggct tgaccccca gcaggtggtg gccatcgcca 960
gcaatggcgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc 1020
aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat ggcggcaagc 1080
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 1140
cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg gagacggtcc 1200
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag gtggtggcca 1260
tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 1320
tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc aataatggtg 1380
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 1440
tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag gcgctggaga 1500
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg 1560
tggccatcgc cagcaatggc ggcggcaggc cggcgctgga gagcattgtg atgc 1614

<210> 280

<211> 5597
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15783

<400> 280

ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggcc aagtctggca tctaccagat taagaacacc 480
ctgaataaca aagtgtatgt tggatctgcc aaggactttg agaaaagatg gaagaggcac 540
ttcaaggatc tggaaaaggg ttgccacagc tccatcaagc tccaaaggtc tttcaacaaa 600
catggcaatg tctttgagtg ctccattttg gaagagatcc cttatgaaaa ggacctcatt 660
atagagaggg aaaacttctg gatcaaggag ctgaactcaa aaatcaatgg atacaacatt 720
gcagatgcca catttggtga cacctgttct actcaccccc ttaaggaaga gatcattaaa 780
aagaggtctg agactgtgaa ggctaagatg ttgaaactgg ggcctgatgg cagaaaggca 840
ctctactcta agccaggtag caaaaatgga aggtggaacc ctgagaccca caagttctgc 900
aaatgtgggg tgagaataca gacatcagcc tacacttgca gcaagtgtag gaacagatca 960
ggagagaata actccttttt caaccacaag cattctcaag gtccatccgg agtggatcta 1020
cgcacgctcg gctacagcca gcagcaacag gagaagatca aaccgaaggt tcgttcgaca 1080
gtggcgcagc accacgaggc actggtcggc cacgggttta cacacgcgca catcgttgcg 1140
ttaagccaac acccggcagc gttagggacc gtcgctgtca agtatcagga catgatcgca 1200
gcgttgccag aggcgacaca cgaagcgatc gttggcgtcg gcaaacagtg gtccggcgca 1260
cgcgctctgg aggccttgct cacggtggcg ggagagttga gaggtccacc gttacagttg 1320
gacacaggcc aacttctcaa gattgcaaaa cgtggcggcg tgaccgcagt ggaggcagtg 1380
catgcatggc gcaatgcact gacgggtgcc ccgctcaact tgaccccgga gcaggtggtg 1440
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 1500
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat 1560
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 1620
ggcttgaccc cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg 1680
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 1740
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 1800
ttgccggtgc tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc 1860
aataatggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 1920
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag 1980
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc 2040
cagcaggtgg tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag 2100
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 2160
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 2220
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 2280
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 2340
acccccccagc aggtggtggc catcgccagc aataatggtg gcaagcaggc gctggagacg 2400
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 2460
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 2520
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 2580
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 2640
ggcttgaccc cccagcaggt ggtggccatc gccagcaata atggtggcaa gcaggcgctg 2700
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 2760
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 2820
ttgccggtgc tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc 2880
aataatggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2940
gcccacggct tgaccccctca gcaggtggtg gccatcgcca gcaatggcgg cggcaggccg 3000
gcgctggaga gcattgttgc ccagttatct cgccctgatc cgtcggccga ctgataagtc 3060
gaccaccacc accaccacca ctgataagag ctcctcgagc ctgcagcagc tgaagctttg 3120
gacttctggc gcgcccaatt cgccctatag tgagtcgtat tacgtcgcgc tcactggccg 3180
tcgttttaca acgtcgtgac tgggaaaacc ctggcgttac ccaacttaat cgccttgcag 3240
cacatccccc tttcgccagc tggcgtaata gcgaagaggc ccgcaccgaa acgcccttcc 3300
caacagttgc gcagcctgaa tggcgaatgg gagcgccctg tagcggcgca ttaagcgcgg 3360
cgggtgtggt ggttacgcgc agcgtgaccg ctacacttgc cagcgcccta gcgcccgctc 3420
ctttcgcttt cttcccttcc tttctcgcca cgttcgccgg ctttccccgt caagctctaa 3480
atcgggggct ccctttaggg ttccgattta gtgctttacg gcacctcgac cccaaaaaac 3540
ttgattaggg tgatggttgg cctgtagtgg gccatagccc tgatagacgg tttttcgccc 3600
tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact 3660
caaccctatc tcggtctatt cttttgattt ataagggatt ttgccgattt cggcctattg 3720
gttaaaaaat gagctgattt aacaaaaatt taacgcgaat tttaacaaaa tattaacgct 3780
tacaatttag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc 3840
taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa 3900
tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt 3960
gcggcatttt gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct 4020
gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc 4080

cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta 4140
tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac 4200
tattctcaga atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc 4260
atgacagtaa gagaattatg cagtgctgcc ataaccatga gtgataacac tgcggccaac 4320
ttacttctga caacgatcgg aggaccgaag gagctaaccg ctttttgca caacatgggg 4380
gatcatgtaa ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac 4440
gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc 4500
gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt 4560
gcaggaccac ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga 4620
gccggtgagc gtggttctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc 4680
cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag 4740
atcgctgaga taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca 4800
tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc 4860
ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca 4920
gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc 4980
tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta 5040
ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt 5100
ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc 5160
gctctgctaa tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg 5220
ttggactcaa gacgatagtt accggataag gcgcagcggt cgggctgaac gggggttcg 5280
tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag 5340
ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc 5400
agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat 5460
agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg 5520
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 5580
tggccttttg ctcacat 5597

<210> 281

<211> 5597
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15784

<400> 281
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggcc aagtctggca tctaccagat taagaacacc 480
ctgaataaca aagtgtatgt tggatctgcc aaggactttg agaaaagatg gaagaggcac 540
ttcaaggatc tggaaaaggg ttgccacagc tccatcaagc tccaaaggtc tttcaacaaa 600
catggcaatg tctttgagtg ctccattttg gaagagatcc cttatgaaaa ggacctcatt 660
atagagaggg aaaacttctg gatcaaggag ctgaactcaa aaatcaatgg atacaacatt 720
gcagatgcca catttggtga cacctgttct actcaccccc ttaaggaaga gatcattaaa 780
aagaggtctg agactgtgaa ggctaagatg ttgaaactgg ggcctgatgg cagaaaggca 840
ctctactcta agccaggtag caaaaatgga aggtggaacc ctgagaccca caagttctgc 900
aaatgtgggg tgagaataca gacatcagcc tacacttgca gcaagtgtag gaacagatca 960
ggagagaata actccttttt caaccacaag cattctcaag gtccatccgg agtggatcta 1020
cgcacgctcg gctacagcca gcagcaacag gagaagatca aaccgaaggt tcgttcgaca 1080
gtggcgcagc accacgaggc actggtcggc cacgggttta cacacgcgca catcgttgcg 1140
ttaagccaac acccggcagc gttagggacc gtcgctgtca agtatcagga catgatcgca 1200
gcgttgccag aggcgacaca cgaagcgatc gttggcgtcg gcaaacagtg gtccggcgca 1260
cgcgctctgg aggccttgct cacggtggcg ggagagttga gaggtccacc gttacagttg 1320
gacacaggcc aacttctcaa gattgcaaaa cgtggcggcg tgaccgcagt ggaggcagtg 1380
catgcatggc gcaatgcact gacgggtgcc ccgctcaact tgacccccca gcaggtggtg 1440
gccatcgcca gcaataatgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg 1500
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagcaatatt 1560
ggtggcaagc aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac 1620
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg 1680
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 1740
gtggtggcca tcgccagcaa taatggtggc aagcaggcgc tggagacggt ccagcggctg 1800

ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 1860
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 1920
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaataatgg tggcaagcag 1980
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgacccca 2040
cagcaggtgg tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag 2100
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 2160
gccagcaata atggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 2220
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 2280
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 2340
accccggagc aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg 2400
gtgcaggcgc tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg 2460
gccatcgcca gcaataatgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg 2520
gtgctgtgcc aggcccacgg cttgacccca cagcaggtgg tggccatcgc cagcaataat 2580
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 2640
ggcttgaccc cccagcaggt ggtggccatc gccagcaata atggtggcaa gcaggcgctg 2700
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag 2760
gtggtggcca tcgccagcaa tattggtggc aagcaggcgc tggagacggt gcaggcgctg 2820
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 2880
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2940
gcccacggct tgacccctca gcaggtggtg gccatcgcca gcaatggcgg cggcaggccg 3000
gcgctggaga gcattgttgc ccagttatct cgccctgatc cgtcggccga ctgataagtc 3060
gaccaccacc accaccacca ctgataagag ctcctcgagc ctgcagcagc tgaagctttg 3120
gacttctggc gcgcccaatt cgccctatag tgagtcgtat tacgtcgcgc tcactggccg 3180
tcgttttaca acgtcgtgac tgggaaaacc ctggcgttac ccaacttaat cgccttgcag 3240
cacatccccc tttcgccagc tggcgtaata gcgaagaggc ccgcaccgaa acgcccttcc 3300
caacagttgc gcagcctgaa tggcgaatgg gagcgccctg tagcggcgca ttaagcgcgg 3360
cgggtgtggt ggttacgcgc agcgtgaccg ctacacttgc cagcgcccta gcgcccgctc 3420
ctttcgcttt cttcccttcc tttctcgcca cgttcgccgg ctttccccgt caagctctaa 3480
atcgggggct ccctttaggg ttccgattta gtgctttacg gcacctcgac cccaaaaaac 3540
ttgattaggg tgatggttgg cctgtagtgg gccatagccc tgatagacgg tttttcgccc 3600
tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact 3660
caaccctatc tcggtctatt cttttgattt ataagggatt ttgccgattt cggcctattg 3720
gttaaaaaat gagctgattt aacaaaaatt taacgcgaat tttaacaaaa tattaacgct 3780
tacaatttag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc 3840
taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa 3900
tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt 3960
gcggcatttt gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct 4020
gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc 4080
cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta 4140
tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac 4200
tattctcaga atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc 4260
atgacagtaa gagaattatg cagtgctgcc ataaccatga gtgataacac tgcggccaac 4320
ttacttctga caacgatcgg aggaccgaag gagctaaccg ctttttgca caacatgggg 4380
gatcatgtaa ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac 4440
gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc 4500
gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt 4560
gcaggaccac ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga 4620
gccggtgagc gtggttctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc 4680
cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag 4740
atcgctgaga taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca 4800
tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc 4860
ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca 4920
gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc 4980
tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta 5040
ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt 5100
ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc 5160
gctctgctaa tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg 5220
ttggactcaa gacgatagtt accggataag gcgcagcggt cgggctgaac gggggttcg 5280
tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag 5340
ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc 5400
agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat 5460
agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg 5520
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 5580
tggccttttg ctcacat 5597

<210> 282

<211> 4440
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS14529

<400> 282
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccccgt tcagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 1860
atcacgaggc cctttcgtct cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg 1920
cagctcccgg agacggtcac agcttgtctg taagcggatg ccgggagcag acaagcccgt 1980
cagggcgcgt cagcgggtgt tggcgggtgt cggggctggc ttaactatgc ggcatcagag 2040
cagattgtac tgagagtgca ccatatgcgg tgtgaaatac cgcacagatg cgtaaggaga 2100
aaataccgca tcaggcgcca ttcgccattc aggctgcgca actgttggga agggcgatcg 2160
gtgcgggcct cttcgctatt acgccagctg gcgaaagggg gatgtgctgc aaggcgatta 2220
agttgggtaa cgccagggtt ttcccagtca cgacgttgta aaacgacggc cagtgccaag 2280
cttgcatgcc tgcagtgcag cgtgaccaaa tctcagaatt ccaatcccac aaaaatctga 2340
gcttaacagc acagttgctc ctctcagagc agaatcgggt attcaacacc ctcatatcaa 2400
ctactacgtt gtgtataacg gtccacatgc cggtatatac gatgactggg gttgtacaaa 2460
ggcggcaaca aacggcgttc ccggagttgc acacaagaaa tttgccacta ttacagaggc 2520
aagagcagca gctgacgcgt acacaacaag tcagcaaaca gacaggttga acttcatccc 2580
caaaggagaa gctcaactca agcccaagag ctttgctaag gccctaacaa gcccaccaaa 2640
gcaaaaagcc cactggctca cgctaggaac caaaaggccc agcagtgatc cagccccaaa 2700
agagatctcc tttgccccgg agattacaat ggacgatttc ctctatcttt acgatctagg 2760
aaggaagttc gaaggtgaag gtgacgacac tatgttcacc actgataatg agaaggttag 2820
cctcttcaat ttcagaaaga atgctgaccc acagatggtt agagaggcct acgcagcagg 2880
tctcatcaag acgatctacc cgagtaacaa tctccaggag atcaaatacc ttcccaagaa 2940
ggttaaagat gcagtcaaaa gattcaggac taattgcatc aagaacacag agaaagacat 3000
atttctcaag atcagaagta ctattccagt atggacgatt caaggcttgc ttcataaacc 3060
aaggcaagta atagagattg gagtctctaa aaaggtagtt cctactgaat ctaaggccat 3120
gcatggagtc taagattcaa atcgaggatc taacagaact cgccgtgaag actggcgaac 3180
agttcataca gagtctttta cgactcaatg acaagaagaa aatcttcgtc aacatggtgg 3240
agcacgacac tctggtctac tccaaaaatg tcaaagatac agtctcagaa gaccaaaggg 3300
ctattgagac ttttcaacaa aggataattt cgggaaacct cctcggattc cattgcccag 3360
ctatctgtca cttcatcgaa aggacagtag aaaaggaagg tggctcctac aaatgccatc 3420
attgcgataa aggaaaggct atcattcaag atctctctgc cgacagtggt cccaaagatg 3480
gacccccacc cacgaggagc atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc 3540
aagtggattg atgtgacatc tccactgacg taagggatga cgcacaatcc cactatcctt 3600

cgcaagaccc ttcctctata taaggaagtt catttcattt ggagaggaca cgcacgagat 3660
ctgcaggtcg actctagagg atcccctcga ggcgcgccac acgcaaacac aaatacacag 3720
cggccttgcc accatggcaa acaccggtgg atcctccgga agatcttcgg ccgactgata 3780
agtcgaccac caccaccacc accactgata agagctcgaa tttccccgat cgttcaaaca 3840
tttggcaata aagtttctta agattgaatc ctgttgccgg tcttgcgatg attatcatat 3900
aatttctgtt gaattacgtt aagcatgtaa taattaacat gtaatgcatg acgttattta 3960
tgagatgggt ttttatgatt agagtcccgc aattatacat ttaatacgcg atagaaaaca 4020
aaatatagcg cgcaaactag gataaattat cgcgcgcggt gtcatctatg ttactagatc 4080
gggaattcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc gctcacaatt 4140
ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc 4200
taactcacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc 4260
cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggcgctct 4320
tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca 4380
gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac 4440

<210> 283

<211> 7005
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS14579

<400> 283
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 1860
atcacgaggc cctttcgtct cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg 1920
cagctcccgg agacggtcac agcttgtctg taagcggatg ccgggagcag acaagcccgt 1980
cagggcgcgt cagcgggtgt tggcgggtgt cggggctggc ttaactatgc ggcatcagag 2040
cagattgtac tgagagtgca ccatatgcgg tgtgaaatac cgcacagatg cgtaaggaga 2100
aaataccgca tcaggcgcca ttcgccattc aggctgcgca actgttggga agggcgatcg 2160
gtgcgggcct cttcgctatt acgccagctg gcgaaagggg gatgtgctgc aaggcgatta 2220
agttgggtaa cgccagggtt ttcccagtca cgacgttgta aaacgacggc cagtgccaag 2280
cttgcatgcc tgcagtgcag cgtgaccaaa tctcagaatt ccaatcccac aaaaatctga 2340
gcttaacagc acagttgctc ctctcagagc agaatcgggt attcaacacc ctcatatcaa 2400
ctactacgtt gtgtataacg gtccacatgc cggtatatac gatgactggg gttgtacaaa 2460
ggcggcaaca aacggcgttc ccggagttgc acacaagaaa tttgccacta ttacagaggc 2520

aagagcagca gctgacgcgt acacaacaag tcagcaaaca gacaggttga acttcatccc 2580
caaaggagaa gctcaactca agcccaagag ctttgctaag gccctaacaa gcccaccaaa 2640
gcaaaaagcc cactggctca cgctaggaac caaaaggccc agcagtgatc cagccccaaa 2700
agagatctcc tttgccccgg agattacaat ggacgatttc ctctatcttt acgatctagg 2760
aaggaagttc gaaggtgaag gtgacgacac tatgttcacc actgataatg agaaggttag 2820
cctcttcaat ttcagaaaga atgctgaccc acagatggtt agagaggcct acgcagcagg 2880
tctcatcaag acgatctacc cgagtaacaa tctccaggag atcaaatacc ttcccaagaa 2940
ggttaaagat gcagtcaaaa gattcaggac taattgcatc aagaacacag agaaagacat 3000
atttctcaag atcagaagta ctattccagt atggacgatt caaggcttgc ttcataaacc 3060
aaggcaagta atagagattg gagtctctaa aaaggtagtt cctactgaat ctaaggccat 3120
gcatggagtc taagattcaa atcgaggatc taacagaact cgccgtgaag actggcgaac 3180
agttcataca gagtctttta cgactcaatg acaagaagaa aatcttcgtc aacatggtgg 3240
agcacgacac tctggtctac tccaaaaatg tcaaagatac agtctcagaa gaccaaaggg 3300
ctattgagac ttttcaacaa aggataattt cgggaaacct cctcggattc cattgcccag 3360
ctatctgtca cttcatcgaa aggacagtag aaaaggaagg tggctcctac aaatgccatc 3420
attgcgataa aggaaaggct atcattcaag atctctctgc cgacagtggt cccaaagatg 3480
gacccccacc cacgaggagc atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc 3540
aagtggattg atgtgacatc tccactgacg taagggatga cgcacaatcc cactatcctt 3600
cgcaagaccc ttcctctata taaggaagtt catttcattt ggagaggaca cgcacgagat 3660
ctgcaggtcg actctagagg atcccctcga ggcgcgccac acgcaaacac aaatacacag 3720
cggccttgcc accatggcca agtctggcat ctaccagatt aagaacaccc tgaataacaa 3780
agtgtatgtt ggatctgcca aggactttga gaaaagatgg aagaggcact tcaaggatct 3840
ggaaaagggt tgccacagct ccatcaagct ccaaaggtct ttcaacaaac atggcaatgt 3900
ctttgagtgc tccattttgg aagagatccc ttatgaaaag gacctcatta tagagaggga 3960
aaacttctgg atcaaggagc tgaactcaaa aatcaatgga tacaacattg cagatgccac 4020
atttggtgac acctgttcta ctcacccct taaggaagag atcattaaaa agaggtctga 4080
gactgtgaag gctaagatgt tgaaactggg gcctgatggc agaaaggcac tctactctaa 4140
gccaggtagc aaaaatggaa ggtggaaccc tgagacccac aagttctgca aatgtggggt 4200
gagaatacag acatcagcct acacttgcag caagtgtagg aacagatcag gagagaataa 4260
ctccttttc aaccacaagc attctcaagg tccatccgga gtggatctac gcacgctcgg 4320
ctacagccag cagcaacagg agaagatcaa accgaaggtt cgttcgacag tggcgcagca 4380
ccacgaggca ctggtcggcc acgggtttac acacgcgcac atcgttgcgt taagccaaca 4440
cccggcagcg ttagggaccg tcgctgtcaa gtatcaggac atgatcgcag cgttgccaga 4500
ggcgacacac gaagcgatcg ttggcgtcgg caaacagtgg tccggcgcac gcgctctgga 4560
ggccttgctc acggtggcgg gagagttgag aggtccaccg ttacagttgg acacaggcca 4620
acttctcaag attgcaaaac gtggcggcgt gaccgcagtg gaggcagtgc atgcatggcg 4680
caatgcactg acgggtgccc cgctcaactt gaccccggag caggtggtgg ccatcgccag 4740
ccacgatggc ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca 4800
ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc agccacgatg gcggcaagca 4860
ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc 4920
ccagcaggtg gtggccatcg ccagcaatgg cggtggcaag caggcgctgg agacggtcca 4980
gcggctgttg ccggtgctgt gccaggccca cggcttgacc ccccagcagg tggtggccat 5040
cgccagcaat ggcggtggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct 5100
gtgccaggcc cacggcttga ccccccagca ggtggtggcc atcgccagca ataatggtgg 5160
caagcaggcg ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt 5220
gaccccggag caggtggtgg ccatcgccag ccacgatggc ggcaagcagg cgctggagac 5280
ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc ttgaccccc agcaggtggt 5340
ggccatcgcc agcaataatg gtggcaagca ggcgctggag acggtccagc ggctgttgcc 5400
ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg gtggccatcg ccagccacga 5460
tggcggcaag caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca 5520
cggcttgacc ccggagcagg tggtggccat cgccagcaat attggtggca agcaggcgct 5580
ggagacggtg caggcgctgt tgccggtgct gtgccaggcc cacggcttga ccccccagca 5640
ggtggtggcc atcgccagca ataatggtgg caagcaggcg ctggagacgg tccagcggct 5700
gttgccggtg ctgtgccagg cccacggctt gaccccggag caggtggtgg ccatcgccag 5760
ccacgatggc ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca 5820
ggcccacggc ttgaccccc agcaggtggt ggccatcgcc agcaatggcg gtggcaagca 5880
ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc 5940
ccagcaggtg gtggccatcg ccagcaataa tggtggcaag caggcgctgg agacggtcca 6000
gcggctgttg ccggtgctgt gccaggccca cggcttgacc ccccagcagg tggtggccat 6060
cgccagcaat ggcggtggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct 6120
gtgccaggcc cacggcttga ccccccagca ggtggtggcc atcgccagca ataatggtgg 6180
caagcaggcg ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt 6240
gacccctcag caggtggtgg ccatcgccag caatggcggc ggcaggccgg cgctggagag 6300
cattgttgcc cagttatctc gccctgatcc gtcggccgac tgataagtcg accaccacca 6360
ccaccaccac tgataagagc tcgaatttcc ccgatcgttc aaacatttgg caataaagtt 6420
tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt ctgttgaatt 6480
acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga tgggttttta 6540
tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata tagcgcgcaa 6600

actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa ttcgtaatca 6660
tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga 6720
gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact cacattaatt 6780
gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga 6840
atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc 6900
actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg 6960
gtaatacggt tatccacaga atcaggggat aacgcaggaa agaac 7005

<210> 284

<211> 7005
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS14581

<400> 284
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
tttttgtttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa atagggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 1860
atcacgaggc cctttcgtct cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg 1920
cagctcccgg agacggtcac agcttgtctg taagcggatg ccgggagcag acaagcccgt 1980
cagggcgcgt cagcgggtgt tggcgggtgt cggggctggc ttaactatgc ggcatcagag 2040
cagattgtac tgagagtgca ccatatgcgg tgtgaaatac cgcacagatg cgtaaggaga 2100
aaataccgca tcaggcgcca ttcgccattc aggctgcgca actgttggga agggcgatcg 2160
gtgcgggcct cttcgctatt acgccagctg gcgaaagggg gatgtgctgc aaggcgatta 2220
agttgggtaa cgccagggtt ttcccagtca cgacgttgta aaacgacggc cagtgccaag 2280
cttgcatgcc tgcagtgcag cgtgaccaaa tctcagaatt ccaatcccac aaaaatctga 2340
gcttaacagc acagttgctc ctctcagagc agaatcgggt attcaacacc ctcatatcaa 2400
ctactacgtt gtgtataacg gtccacatgc cggtatatac gatgactggg gttgtacaaa 2460
ggcggcaaca aacggcgttc ccggagttgc acacaagaaa tttgccacta ttacagaggc 2520
aagagcagca gctgacgcgt acacaacaag tcagcaaaca gacaggttga acttcatccc 2580
caaaggagaa gctcaactca agcccaagag ctttgctaag gccctaacaa gcccaccaaa 2640
gcaaaaagcc cactggctca cgctaggaac caaaaggccc agcagtgatc cagccccaaa 2700
agagatctcc tttgccccgg agattacaat ggacgatttc ctctatcttt acgatctagg 2760
aaggaagttc gaaggtgaag gtgacgacac tatgttcacc actgataatg agaaggttag 2820
cctcttcaat ttcagaaaga atgctgaccc acagatggtt agagaggcct acgcagcagg 2880
tctcatcaag acgatctacc cgagtaacaa tctccaggag atcaaatacc ttcccaagaa 2940

ggttaaagat gcagtcaaaa gattcaggac taattgcatc aagaacacag agaaagacat 3000
atttctcaag atcagaagta ctattccagt atggacgatt caaggcttgc ttcataaacc 3060
aaggcaagta atagagattg gagtctctaa aaaggtagtt cctactgaat ctaaggccat 3120
gcatggagtc taagattcaa atcgaggatc taacagaact cgccgtgaag actggcgaac 3180
agttcataca gagtctttta cgactcaatg acaagaagaa aatcttcgtc aacatggtgg 3240
agcacgacac tctggtctac tccaaaaatg tcaaagatac agtctcagaa gaccaaaggg 3300
ctattgagac ttttcaacaa aggataattt cgggaaacct cctcggattc cattgcccag 3360
ctatctgtca cttcatcgaa aggacagtag aaaaggaagg tggctcctac aaatgccatc 3420
attgcgataa aggaaaggct atcattcaag atctctctgc cgacagtggt cccaaagatg 3480
gacccccacc cacgaggagc atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc 3540
aagtggattg atgtgacatc tccactgacg taagggatga cgcacaatcc cactatcctt 3600
cgcaagaccc ttcctctata taaggaagtt catttcattt ggagaggaca cgcacgagat 3660
ctgcaggtcg actctagagg atcccctcga ggcgcgccac acgcaaacac aaatacacag 3720
cggccttgcc accatggcca agtctggcat ctaccagatt aagaacaccc tgaataacaa 3780
agtgtatgtt ggatctgcca aggactttga gaaaagatgg aagaggcact tcaaggatct 3840
ggaaaagggt tgccacagct ccatcaagct ccaaaggtct ttcaacaaac atggcaatgt 3900
ctttgagtgc tccattttgg aagagatccc ttatgaaaag gacctcatta tagagaggga 3960
aaacttctgg atcaaggagc tgaactcaaa aatcaatgga tacaacattg cagatgccac 4020
atttggtgac acctgttcta ctcacccct taaggaagag atcattaaaa agaggtctga 4080
gactgtgaag gctaagatgt tgaaactggg gcctgatggc agaaaggcac tctactctaa 4140
gccaggtagc aaaaatggaa ggtggaaccc tgagacccac aagttctgca aatgtggggt 4200
gagaatacag acatcagcct acacttgcag caagtgtagg aacagatcag gagagaataa 4260
ctccttttc aaccacaagc attctcaagg tccatccgga gtggatctac gcacgctcgg 4320
ctacagccag cagcaacagg agaagatcaa accgaaggtt cgttcgacag tggcgcagca 4380
ccacgaggca ctggtcggcc acgggtttac acacgcgcac atcgttgcgt taagccaaca 4440
cccggcagcg ttagggaccg tcgctgtcaa gtatcaggac atgatcgcag cgttgccaga 4500
ggcgacacac gaagcgatcg ttggcgtcgg caaacagtgg tccggcgcac gcgctctgga 4560
ggccttgctc acggtggcgg gagagttgag aggtccaccg ttacagttgg acacaggcca 4620
acttctcaag attgcaaaac gtggcggcgt gaccgcagtg gaggcagtgc atgcatggcg 4680
caatgcactg acgggtgccc cgctcaactt gacccccag caggtggtgg ccatcgccag 4740
caataatggt ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca 4800
ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc agcaatattg gtggcaagca 4860
ggcgctggag acggtgcagg cgctgttgcc ggtgctgtgc caggcccacg gcttgacccc 4920
ggagcaggtg gtggccatcg ccagcaatat tggtggcaag caggcgctgg agacggtgca 4980
ggcgctgttg ccggtgctgt gccaggccca cggcttgacc ccccagcagg tggtggccat 5040
cgccagcaat aatggtggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct 5100
gtgccaggcc cacggcttga ccccggagca ggtggtggcc atcgccagcc acgatggcgg 5160
caagcaggcg ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt 5220
gacccccag caggtggtgg ccatcgccag caataatggt ggcaagcagg cgctggagac 5280
ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc ttgacccccc agcaggtggt 5340
ggccatcgcc agcaataatg gtggcaagca ggcgctggag acggtccagc ggctgttgcc 5400
ggtgctgtgc caggcccacg gcttgacccc ccagcaggtg gtggccatcg ccagcaataa 5460
tggtggcaag caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca 5520
cggcttgacc ccggagcagg tggtggccat cgccagcaat attggtggca agcaggcgct 5580
ggagacggtg caggcgctgt tgccggtgct gtgccaggcc cacggcttga ccccggagca 5640
ggtggtggcc atcgccagca atattggtgg caagcaggcg ctggagacgg tgcaggcgct 5700
gttgccggtg ctgtgccagg cccacggctt gacccccag caggtggtgg ccatcgccag 5760
caataatggt ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca 5820
ggcccacggc ttgacccccc agcaggtggt ggccatcgcc agcaataatg gtggcaagca 5880
ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc 5940
ccagcaggtg gtggccatcg ccagcaataa tggtggcaag caggcgctgg agacggtcca 6000
gcggctgttg ccggtgctgt gccaggccca cggcttgacc ccggagcagg tggtggccat 6060
cgccagcaat attggtggca agcaggcgct ggagacggtg caggcgctgt tgccggtgct 6120
gtgccaggcc cacggcttga ccccggagca ggtggtggcc atcgccagcc acgatggcgg 6180
caagcaggcg ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt 6240
gacccctcag caggtggtgg ccatcgccag caatggcggc ggcaggccgg cgctggagag 6300
cattgttgcc cagttatctc gccctgatcc gtcggccgac tgataagtcg accaccacca 6360
ccaccaccac tgataagagc tcgaatttcc ccgatcgttc aaacatttgg caataaagtt 6420
tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt ctgttgaatt 6480
acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga tgggttttta 6540
tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata tagcgcgcaa 6600
actaggataa attatcgcgc gcggtgtcat ctatgttact agatcgggaa ttcgtaatca 6660
tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga 6720
gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact cacattaatt 6780
gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga 6840
atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc 6900
actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg 6960
gtaatacggt tatccacaga atcaggggat aacgcaggaa agaac 7005

<210> 285

<211> 4208
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15785

<400> 285
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aaggaggaaa cattcatgga 420
actgaagaac tccatcagcg attacaccga agccgagttc gtgcagctgt tgaaagaaat 480
tgagaaagag aacgttgctg ccaccgatga cgtgctggat gttttgctgg aacattttgt 540
gaagatcact gaacatccag atggcactga cctgatctac tacccgagcg ataaccgtga 600
tgacagccca gaaggcattg ttaaagagat caaagaatgg agagctgcca acggcaagcc 660
aggcttcaaa cagggttgac ggccaactga tttaattaag cgcgccacac gcaaacacaa 720
atacacagcg gccttgccac catggtggat ctacgcacgc tcggctacag ccagcagcaa 780
caggagaaga tcaaaccgaa ggttcgttcg acagtggcgc agcaccacga ggcactggtc 840
ggccacgggt ttacacacgc gcacatcgtt gcgttaagcc aacacccggc agcgttaggg 900
accgtcgctg tcaagtatca ggacatgatc gcagcgttgc cagaggcgac acacgaagcg 960
atcgttggcg tcggcaaaca gtggtccggc gcacgcgctc tggaggcctt gctcacggtg 1020
gcgggagagt tgagaggtcc accgttacag ttggacacag gccaacttct caagattgca 1080
aaacgtggcg gcgtgaccgc agtggaggca gtgcatgcat ggcgcaatgc actgacgggt 1140
gccccgctca acttgaccgg agacgcccgg gggatcaggt cacgtgcgtc tcggagcatt 1200
gttgcccagt tatctcgccc tgatccggga tcctccggat ctaaacgcaa caaaccaggt 1260
aaggccactg gcaaaggtaa accagtgaac aataagtggc tgaacaacgc tggcaaagac 1320
ttgggttctc cagtgccgga tcgcattgcc aacaagctga gagacaaaga attcaagagt 1380
tttgatgact ttcgcgcaaa attctgggaa gaagtgtcta aagatccaga actgtctaaa 1440
cagtttagtc gcaacaacaa tgacagaatg aaagttggta aagctccgaa gacccgcacc 1500
caagacgtgt ctggcaaacg caccagtttt gaattgcatc atgaaaaacc aatctctcag 1560
aatggtggcg tttacgacat ggacaacatc tctgtggtta ctccgaaaag acatatcgat 1620
attcatcgcg gtaaaggatc caccggtgac tgatcggccg actgataagt cgaccaccac 1680
caccaccacc actgataaga gctcctcgag cctgcagcag ctgaagcttt ggacttctgg 1740
cgcgcccaat tcgccctata gtgagtcgta ttacgtcgcg ctcactggcc gtcgttttac 1800
aacgtcgtga ctgggaaaac cctggcgtta cccaacttaa tcgccttgca gcacatcccc 1860
ctttcgccag ctggcgtaat agcgaagagg cccgcaccga aacgcccttc ccaacagttg 1920
cgcagcctga atggcgaatg ggagcgccct gtagcggcgc attaagcgcg gcgggtgtgg 1980
tggttacgcg cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt 2040
tcttcccttc ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc 2100
tccctttagg gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg 2160
gtgatggttg gcctgtagtg ggccatagcc ctgatagacg gtttttcgcc ctttgacgtt 2220
ggagtccacg ttctttaata gtggactctt gttccaaact ggaacaacac tcaaccctat 2280
ctcggtctat tcttttgatt tataagggat tttgccgatt tcggcctatt ggttaaaaaa 2340
tgagctgatt taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc ttacaattta 2400
ggtggcactt ttcggggaaa tgtgcgcgga acccctattt gtttatttttt ctaaatacat 2460
tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa 2520
aggaagagta tgagtattca acatttccgt gtcgccctta ttccctttttt tgcggcattt 2580
tgccttcctg tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag 2640
ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaagat ccttgagagt 2700
tttcgccccg aagaacgttt tccaatgatg agcacttttа aagttctgct atgtggcgcg 2760
gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag 2820
aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta 2880
agagaattat gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg 2940
acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta 3000
actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac 3060
accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt 3120
actctagctt cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca 3180
cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag 3240
cgtggttctc gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta 3300
gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag 3360

ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt 3420
tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat ccttttttgat 3480
aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta 3540
gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa 3600
acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt 3660
tttccgaagg taactggctt cagcagagcg cagataccaa atactgttct tctagtgtag 3720
ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta 3780
atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca 3840
agacgatagt taccggataa ggcgcagcgg tcgggctgaa cgggggttc gtgcacacag 3900
cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa 3960
agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga 4020
acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc 4080
gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc 4140
ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt 4200
gctcacat 4208

<210> 286

<211> 5756
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15786

<400> 286
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aaggaggaaa cattcatgga 420
actgaagaac tccatcagcg attacaccga agccgagttc gtgcagctgt tgaaagaaat 480
tgagaaagag aacgttgctg ccaccgatga cgtgctggat gttttgctgg aacattttgt 540
gaagatcact gaacatccag atggcactga cctgatctac tacccgagcg ataaccgtga 600
tgacagccca gaaggcattg ttaaagagat caaagaatgg agagctgcca acggcaagcc 660
aggcttcaaa cagggttgac ggccaactga tttaattaag cgcgccacac gcaaacacaa 720
atacacagcg gccttgccac catggtggat ctacgcacgc tcggctacag ccagcagcaa 780
caggagaaga tcaaaccgaa ggttcgttcg acagtggcgc agcaccacga ggcactggtc 840
ggccacgggt ttacacacgc gcacatcgtt gcgttaagcc aacacccggc agcgttaggg 900
accgtcgctg tcaagtatca ggacatgatc gcagcgttgc cagaggcgac acacgaagcg 960
atcgttggcg tcggcaaaca gtggtccggc gcacgcgctc tggaggcctt gctcacggtg 1020
gcgggagagt tgagaggtcc accgttacag ttggacacag gccaacttct caagattgca 1080
aaacgtggcg gcgtgaccgc agtggaggca gtgcatgcat ggcgcaatgc actgacgggt 1140
gccccgctca acttgacccc ggagcaggtg gtggccatcg ccagccacga tggcggcaag 1200
caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc 1260
ccggagcagg tggtggccat cgccagccac gatggcggca agcaggcgct ggagacggtc 1320
cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccccagca ggtggtggcc 1380
atcgccagca atggcggtgg caagcaggcg ctggagacgg tccagcggct gttgccggtg 1440
ctgtgccagg cccacggctt gaccccccag caggtggtgg ccatcgccag caatggcggt 1500
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc 1560
ttgacccccc agcaggtggt ggccatcgcc agcaataatg gtggcaagca ggcgctggag 1620
acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg 1680
gtggccatcg ccagccacga tggcggcaag caggcgctgg agacggtcca gcggctgttg 1740
ccggtgctgt gccaggccca cggcttgacc ccccagcagg tggtggccat cgccagcaat 1800
aatggtggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc 1860
cacggcttga ccccggagca ggtggtggcc atcgccagcc acgatggcgg caagcaggcg 1920
ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt gaccccggag 1980
caggtggtgg ccatcgccag caatattggt ggcaagcagg cgctggagac ggtgcaggcg 2040
ctgttgccgg tgctgtgcca ggcccacggc ttgacccccc agcaggtggt ggccatcgcc 2100
agcaataatg gtggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc 2160
caggcccacg gcttgacccc ggagcaggtg gtggccatcg ccagccacga tggcggcaag 2220
caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc 2280
ccccagcagg tggtggccat cgccagcaat ggcggtggca agcaggcgct ggagacggtc 2340
cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccccagca ggtggtggcc 2400
atcgccagca ataatggtgg caagcaggcg ctggagacgg tccagcggct gttgccggtg 2460

ctgtgccagg cccacggctt gacccccag caggtggtgg ccatcgccag caatggcggt 2520
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc 2580
ttgacccccc agcaggtggt ggccatcgcc agcaataatg gtggcaagca ggcgctggag 2640
acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc tcagcaggtg 2700
gtggccatcg ccagcaatgg cggcggcagg ccggcgctgg agagcattgt tgcccagtta 2760
tctcgccctg atccgggatc ctccggatct aaacgcaaca aaccaggtaa ggccactggc 2820
aaaggtaaac cagtgaacaa taagtggctg aacaacgctg gcaaagactt gggttctcca 2880
gtgccggatc gcattgccaa caagctgaga gacaaagaat tcaagagttt tgatgacttt 2940
cgcgcaaaat tctgggaaga agtgtctaaa gatccagaac tgtctaaaca gtttagtcgc 3000
aacaacaatg acagaatgaa agttggtaaa gctccgaaga cccgcaccca agacgtgtct 3060
ggcaaacgca ccagttttga attgcatcat gaaaaaccaa tctctcagaa tggtggcgtt 3120
tacgacatgg acaacatctc tgtggttact ccgaaaagac atatcgatat tcatcgcggt 3180
aaaggatcca ccggtgactg atcggccgac tgataagtcg accaccacca ccaccaccac 3240
tgataagagc tcctcgagcc tgcagcagct gaagctttgg acttctggcg cgcccaattc 3300
gccctatagt gagtcgtatt acgtcgcgct cactggccgt cgttttacaa cgtcgtgact 3360
gggaaaaccc tggcgttacc caacttaatc gccttgcagc acatccccct ttcgccagct 3420
ggcgtaatag cgaagaggcc cgcaccgaaa cgcccttccc aacagttgcg cagcctgaat 3480
ggcgaatggg agcgccctgt agcggcgcat taagcgcggc gggtgtggtg gttacgcgca 3540
gcgtgaccgc tacacttgcc agcgccctag cgcccgctcc tttcgctttc ttcccttcct 3600
ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcgggggctc cctttagggt 3660
tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggt gatggttggc 3720
ctgtagtggg ccatagccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 3780
ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc 3840
ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 3900
acaaaaattt aacgcgaatt ttaacaaaat attaacgctt acaatttagg tggcactttt 3960
cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc aaatatgtat 4020
ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag gaagagtatg 4080
agtattcaac atttccgtgt cgcccttatt cccttttttg cggcattttg ccttcctgtt 4140
tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt gggtgcacga 4200
gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa 4260
gaacgttttc caatgatgag cacttttaaa gttctgctat gtggcgcggt attatcccgt 4320
attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt 4380
gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc 4440
agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac aacgatcgga 4500
ggaccgaagg agctaaccgc ttttttgcac aacatggggg atcatgtaac tcgccttgat 4560
cgttgggaac cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct 4620
gtagcaatgg caacaacgtt gcgcaaacta ttaactggcg aactacttac tctagcttcc 4680
cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg 4740
gcccttccgg ctggctggtt tattgctgat aaatctggag ccggtgagcg tggttctcgc 4800
ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg 4860
acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca 4920
ctgattaagc attggtaact gtcagaccaa gtttactcat atatacttta gattgattta 4980
aaacttcatt tttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc 5040
aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa 5100
ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca 5160
ccgctaccag cggtggtttg tttgccggat caagagctac caactctttt tccgaaggta 5220
actggcttca gcagagcgca gataccaaat actgttcttc tagtgtagcc gtagttaggc 5280
caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaat cctgttacca 5340
gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta 5400
ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttggag 5460
cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag cgccacgctt 5520
cccgaaggga gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc 5580
acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac 5640
ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac 5700
gccagcaacg cggccttttt acggttcctg gcctttgct ggccttttgc tcacat 5756

<210> 287

<211> 5756
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15787

<400> 287
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60

tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aaggaggaaa cattcatgga 420
actgaagaac tccatcagcg attacaccga agccgagttc gtgcagctgt tgaaagaaat 480
tgagaaagag aacgttgctg ccaccgatga cgtgctggat gttttgctgg aacattttgt 540
gaagatcact gaacatccag atggcactga cctgatctac tacccgagcg ataaccgtga 600
tgacagccca gaaggcattg ttaaagagat caaagaatgg agagctgcca acggcaagcc 660
aggcttcaaa cagggttgac ggccaactga tttaattaag cgcgccacac gcaaacacaa 720
atacacagcg gccttgccac catggtggat ctacgcacgc tcggctacag ccagcagcaa 780
caggagaaga tcaaaccgaa ggttcgttcg acagtggcgc agcaccacga ggcactggtc 840
ggccacgggt ttacacacgc gcacatcgtt gcgttaagcc aacacccggc agcgttaggg 900
accgtcgctg tcaagtatca ggacatgatc gcagcgttgc cagaggcgac acacgaagcg 960
atcgttggcg tcggcaaaca gtggtccggc gcacgcgctc tggaggcctt gctcacggtg 1020
gcgggagagt tgagaggtcc accgttacag ttggacacag gccaacttct caagattgca 1080
aaacgtggcg gcgtgaccgc agtggaggca gtgcatgcat ggcgcaatgc actgacgggt 1140
gccccgctca acttgacccc ccagcaggtg gtggccatcg ccagcaataa tggtggcaag 1200
caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc 1260
ccggagcagg tggtggccat cgccagcaat attggtggca agcaggcgct ggagacggtg 1320
caggcgctgt tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc 1380
atcgccagca atattggtgg caagcaggcg ctggagacgg tgcaggcgct gttgccggtg 1440
ctgtgccagg cccacggctt gacccccag caggtggtgg ccatcgccag caataatggt 1500
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc 1560
ttgaccccgg agcaggtggt ggccatcgcc agccacgatg gcggcaagca ggcgctggag 1620
acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc ccagcaggtg 1680
gtggccatcg ccagcaataa tggtggcaag caggcgctgg agacggtcca gcggctgttg 1740
ccggtgctgt gccaggccca cggcttgacc ccccagcagg tggtggccat cgccagcaat 1800
aatggtggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc 1860
cacggcttga ccccccagca ggtggtggcc atcgccagca ataatggtgg caagcaggcg 1920
ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt gaccccggag 1980
caggtggtgg ccatcgccag caatattggt ggcaagcagg cgctggagac ggtgcaggcg 2040
ctgttgccgg tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc 2100
agcaatattg gtggcaagca ggcgctggag acggtgcagg cgctgttgcc ggtgctgtgc 2160
caggcccacg gcttgacccc ccagcaggtg gtggccatcg ccagcaataa tggtggcaag 2220
caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc 2280
ccccagcagg tggtggccat cgccagcaat aatggtggca agcaggcgct ggagacggtc 2340
cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccccagca ggtggtggcc 2400
atcgccagca ataatggtgg caagcaggcg ctggagacgg tccagcggct gttgccggtg 2460
ctgtgccagg cccacggctt gaccccggag caggtggtgg ccatcgccag caatattggt 2520
ggcaagcagg cgctggagac ggtgcaggcg ctgttgccgg tgctgtgcca ggcccacggc 2580
ttgaccccgg agcaggtggt ggccatcgcc agccacgatg gcggcaagca ggcgctggag 2640
acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc tcagcaggtg 2700
gtggccatcg ccagcaatgg cggcggcagg ccggcgctgg agagcattgt tgcccagtta 2760
tctcgccctg atccgggatc ctccggatct aaacgcaaca aaccaggtaa ggccactggc 2820
aaaggtaaac cagtgaacaa taagtggctg aacaacgctg gcaaagactt gggttctcca 2880
gtgccggatc gcattgccaa caagctgaga gacaaagaat tcaagagttt tgatgacttt 2940
cgcgcaaaat tctgggaaga agtgtctaaa gatccagaac tgtctaaaca gtttagtcgc 3000
aacaacaatg acagaatgaa agttggtaaa gctccgaaga cccgcaccca agacgtgtct 3060
ggcaaacgca ccagttttga attgcatcat gaaaaaccaa tctctcagaa tggtggcgtt 3120
tacgacatgg acaacatctc tgtggttact ccgaaaagac atatcgatat tcatcgcggt 3180
aaaggatcca ccggtgactg atcggccgac tgataagtcg accaccacca ccaccaccac 3240
tgataagagc tcctcgagcc tgcagcagct gaagctttgg acttctggcg cgcccaattc 3300
gccctatagt gagtcgtatt acgtcgcgct cactggccgt cgttttacaa cgtcgtgact 3360
gggaaaaccc tggcgttacc caacttaatc gccttgcagc acatccccct ttcgccagct 3420
ggcgtaatag cgaagaggcc cgcaccgaaa cgcccttccc aacagttgcg cagcctgaat 3480
ggcgaatggg agcgccctgt agcggcgcat taagcgcggc gggtgtggtg gttacgcgca 3540
gcgtgaccgc tacacttgcc agcgccctag cgcccgctcc tttcgctttc ttcccttcct 3600
ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcgggggctc cctttagggt 3660
tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggt gatggttggc 3720
ctgtagtggg ccatagccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 3780
ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc 3840
ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 3900
acaaaaattt aacgcgaatt ttaacaaaat attaacgctt acaatttagg tggcactttt 3960
cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc aaatatgtat 4020
ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag gaagagtatg 4080
agtattcaac atttccgtgt cgcccttatt ccctttttttg cggcattttg ccttcctgtt 4140

tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt gggtgcacga 4200
gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa 4260
gaacgttttc caatgatgag cacttttaaa gttctgctat gtggcgcggt attatcccgt 4320
attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt 4380
gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc 4440
agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac aacgatcgga 4500
ggaccgaagg agctaaccgc ttttttgcac aacatggggg atcatgtaac tcgccttgat 4560
cgttgggaac cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct 4620
gtagcaatgg caacaacgtt gcgcaaacta ttaactggcg aactacttac tctagcttcc 4680
cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg 4740
gcccttccgg ctggctggtt tattgctgat aaatctggag ccggtgagcg tggttctcgc 4800
ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg 4860
acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca 4920
ctgattaagc attggtaact gtcagaccaa gtttactcat atatacttta gattgattta 4980
aaacttcatt tttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc 5040
aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa 5100
ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca 5160
ccgctaccag cggtggtttg tttgccggat caagagctac caactctttt tccgaaggta 5220
actggcttca gcagagcgca gataccaaat actgttcttc tagtgtagcc gtagttaggc 5280
caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaat cctgttacca 5340
gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta 5400
ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttggag 5460
cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag cgccacgctt 5520
cccgaaggga gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc 5580
acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac 5640
ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac 5700
gccagcaacg cggccttttt acggttcctg gccttttgct ggccttttgc tcacat 5756

<210> 288

<211> 6867
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS14584

<400> 288
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740

atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 1860
atcacgaggc cctttcgtct cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg 1920
cagctcccgg agacggtcac agcttgtctg taagcggatg ccgggagcag acaagcccgt 1980
cagggcgcgt cagcgggtgt tggcgggtgt cggggctggc ttaactatgc ggcatcagag 2040
cagattgtac tgagagtgca ccatatgcgg tgtgaaatac cgcacagatg cgtaaggaga 2100
aaataccgca tcaggcgcca ttcgccattc aggctgcgca actgttggga agggcgatcg 2160
gtgcgggcct cttcgctatt acgccagctg gcgaaagggg gatgtgctgc aaggcgatta 2220
agttgggtaa cgccagggtt ttcccagtca cgacgttgta aaacgacggc cagtgccaag 2280
cttgcatgcc tgcagtgcag cgtgaccaaa tctcagaatt ccaatcccac aaaaatctga 2340
gcttaacagc acagttgctc ctctcagagc agaatcgggt attcaacacc ctcatatcaa 2400
ctactacgtt gtgtataacg gtccacatgc cggtatatac gatgactggg gttgtacaaa 2460
ggcggcaaca aacggcgttc ccggagttgc acacaagaaa tttgccacta ttacagaggc 2520
aagagcagca gctgacgcgt acacaacaag tcagcaaaca gacaggttga acttcatccc 2580
caaaggagaa gctcaactca agcccaagag ctttgctaag gccctaacaa gcccaccaaa 2640
gcaaaaagcc cactggctca cgctaggaac caaaaggccc agcagtgatc cagccccaaa 2700
agagatctcc tttgccccgg agattacaat ggacgatttc ctctatcttt acgatctagg 2760
aaggaagttc gaaggtgaag gtgacgacac tatgttcacc actgataatg agaaggttag 2820
cctcttcaat ttcagaaaga atgctgaccc acagatggtt agagaggcct acgcagcagg 2880
tctcatcaag acgatctacc cgagtaacaa tctccaggag atcaaatacc ttcccaagaa 2940
ggttaaagat gcagtcaaaa gattcaggac taattgcatc aagaacacag agaaagacat 3000
atttctcaag atcagaagta ctattccagt atggacgatt caaggcttgc ttcataaacc 3060
aaggcaagta atagagattg gagtctctaa aaaggtagtt cctactgaat ctaaggccat 3120
gcatggagtc taagattcaa atcgaggatc taacagaact cgccgtgaag actggcgaac 3180
agttcataca gagtctttta cgactcaatg acaagaagaa aatcttcgtc aacatggtgg 3240
agcacgacac tctggtctac tccaaaaatg tcaaagatac agtctcagaa gaccaaaggg 3300
ctattgagac ttttcaacaa aggataattt cgggaaacct cctcggattc cattgcccag 3360
ctatctgtca cttcatcgaa aggacagtag aaaaggaagg tggctcctac aaatgccatc 3420
attgcgataa aggaaaggct atcattcaag atctctctgc cgacagtggt cccaaagatg 3480
gacccccacc cacgaggagc atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc 3540
aagtggattg atgtgacatc tccactgacg taagggatga cgcacaatcc cactatcctt 3600
cgcaagaccc ttcctctata taaggaagtt catttcattt ggagaggaca cgcacgagat 3660
ctgcaggtcg actctagagg atcccctcga ggcgcgccac acgcaaacac aaatacacag 3720
cggccttgcc accatggtgg atctacgcac gctcggctac agccagcagc aacaggagaa 3780
gatcaaaccg aaggttcgtt cgacagtggc gcagcaccac gaggcactgg tcggccacgg 3840
gtttacacac gcgcacatcg ttgcgttaag ccaacacccg gcagcgttag ggaccgtcgc 3900
tgtcaagtat caggacatga tcgcagcgtt gccagaggcg acacacgaag cgatcgttgg 3960
cgtcggcaaa cagtggtccg gcgcacgcgc tctggaggcc ttgctcacgg tggcgggaga 4020
gttgagaggt ccaccgttac agttggacac aggccaactt ctcaagattg caaaacgtgg 4080
cggcgtgacc gcagtggagg cagtgcatgc atggcgcaat gcactgacgg gtgccccgct 4140
caacttgacc ccggagcagg tggtggccat cgccagccac gatggcggca agcaggcgct 4200
ggagacggtc cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccggagca 4260
ggtggtggcc atcgccagcc acgatggcgg caagcaggcg ctggagacgg tccagcggct 4320
gttgccggtg ctgtgccagg cccacggctt gacccccag caggtggtgg ccatcgccag 4380
caatggcggt ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca 4440
ggcccacggc ttgacccccc agcaggtggt ggccatcgcc agcaatggcg gtggcaagca 4500
ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc 4560
ccagcaggtg gtggccatcg ccagcaataa tggtggcaag caggcgctgg agacggtcca 4620
gcggctgttg ccggtgctgt gccaggccca cggcttgacc ccggagcagg tggtggccat 4680
cgccagccac gatggcggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct 4740
gtgccaggcc cacggcttga ccccccagca ggtggtggcc atcgccagca ataatggtgg 4800
caagcaggcg ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt 4860
gaccccggag caggtggtgg ccatcgccag ccacgatggc ggcaagcagg cgctggagac 4920
ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt 4980
ggccatcgcc agcaatattg gtggcaagca ggcgctggag acggtgcagg cgctgttgcc 5040
ggtgctgtgc caggcccacg gcttgacccc ccagcaggtg gtggccatcg ccagcaataa 5100
tggtggcaag caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca 5160
cggcttgacc ccggagcagg tggtggccat cgccagccac gatggcggca agcaggcgct 5220
ggagacggtc cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccccagca 5280
ggtggtggcc atcgccagca atggcggtgg caagcaggcg ctggagacgg tccagcggct 5340
gttgccggtg ctgtgccagg cccacggctt gacccccag caggtggtgg ccatcgccag 5400
caataatggt ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca 5460
ggcccacggc ttgacccccc agcaggtggt ggccatcgcc agcaatggcg gtggcaagca 5520
ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc 5580
ccagcaggtg gtggccatcg ccagcaataa tggtggcaag caggcgctgg agacggtcca 5640
gcggctgttg ccggtgctgt gccaggccca cggcttgacc cctcagcagg tggtggccat 5700
cgccagcaat ggcggcggca ggccggcgct ggagagcatt gttgcccagt tatctcgccc 5760
tgatccggga tcctccggat ctaaacgcaa caaaccaggt aaggccactg gcaaaggtaa 5820

accagtgaac aataagtggc tgaacaacgc tggcaaagac ttgggttctc cagtgccgga 5880
tcgcattgcc aacaagctga gagacaaaga attcaagagt tttgatgact ttcgcgcaaa 5940
attctgggaa gaagtgtcta aagatccaga actgtctaaa cagtttagtc gcaacaacaa 6000
tgacagaatg aaagttggta aagctccgaa gacccgcacc caagacgtgt ctggcaaacg 6060
caccagtttt gaattgcatc atgaaaaacc aatctctcag aatggtggcg tttacgacat 6120
ggacaacatc tctgtggtta ctccgaaaag acatatcgat attcatcgcg gtaaaggatc 6180
caccggtgac tgatcggccg actgataagt cgaccaccac caccaccacc actgataaga 6240
gctcgaattt ccccgatcgt tcaaacattt ggcaataaag tttcttaaga ttgaatcctg 6300
ttgccggtct tgcgatgatt atcatataat ttctgttgaa ttacgttaag catgtaataa 6360
ttaacatgta atgcatgacg ttatttatga gatgggtttt tatgattaga gtcccgcaat 6420
tatacattta atacgcgata gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc 6480
gcgcggtgtc atctatgtta ctagatcggg aattcgtaat catggtcata gctgtttcct 6540
gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag cataaagtgt 6600
aaagcctggg gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc 6660
gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg 6720
agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg 6780
gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca 6840
gaatcagggg ataacgcagg aaagaac 6867

<210> 289

<211> 6867
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS14587

<400> 289
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt 1860
atcacgaggc cctttcgtct cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg 1920
cagctcccgg agacggtcac agcttgtctg taagcggatg ccgggagcag acaagcccgt 1980
cagggcgcgt cagcgggtgt tggcgggtgt cggggctggc ttaactatgc ggcatcagag 2040
cagattgtac tgagagtgca ccatatgcgg tgtgaaatac cgcacagatg cgtaaggaga 2100
aaataccgca tcaggcgcca ttcgccattc aggctgcgca actgttggga agggcgatcg 2160
gtgcgggcct cttcgctatt acgccagctg gcgaaagggg gatgtgctgc aaggcgatta 2220
agttgggtaa cgccagggtt ttcccagtca cgacgttgta aaacgacggc cagtgccaag 2280

```
cttgcatgcc tgcagtgcag cgtgaccaaa tctcagaatt ccaatcccac aaaaatctga   2340
gcttaacagc acagttgctc ctctcagagc agaatcgggt attcaacacc ctcatatcaa   2400
ctactacgtt gtgtataacg gtccacatgc cggtatatac gatgactggg gttgtacaaa   2460
ggcggcaaca aacggcgttc ccggagttgc acacaagaaa tttgccacta ttacagaggc   2520
aagagcagca gctgacgcgt acacaacaag tcagcaaaca gacaggttga acttcatccc   2580
caaaggagaa gctcaactca agcccaagag ctttgctaag gccctaacaa gcccaccaaa   2640
gcaaaaagcc cactggctca cgctaggaac caaaaggccc agcagtgatc cagccccaaa   2700
agagatctcc tttgccccgg agattacaat ggacgatttc ctctatcttt acgatctagg   2760
aaggaagttc gaaggtgaag gtgacgacac tatgttcacc actgataatg agaaggttag   2820
cctcttcaat ttcagaaaga atgctgaccc acagatggtt agagaggcct acgcagcagg   2880
tctcatcaag acgatctacc cgagtaacaa tctccaggag atcaaatacc ttcccaagaa   2940
ggttaaagat gcagtcaaaa gattcaggac taattgcatc aagaacacag agaaagacat   3000
atttctcaag atcagaagta ctattccagt atggacgatt caaggcttgc ttcataaacc   3060
aaggcaagta atagagattg gagtctctaa aaaggtagtt cctactgaat ctaaggccat   3120
gcatggagtc taagattcaa atcgaggatc taacagaact cgccgtgaag actggcgaac   3180
agttcataca gagtctttta cgactcaatg acaagaagaa aatcttcgtc aacatggtgg   3240
agcacgacac tctggtctac tccaaaaatg tcaaagatac agtctcagaa gaccaaaggg   3300
ctattgagac ttttcaacaa aggataattt cgggaaacct cctcggattc cattgcccag   3360
ctatctgtca cttcatcgaa aggacagtag aaaaggaagg tggctcctac aaatgccatc   3420
attgcgataa aggaaaggct atcattcaag atctctctgc cgacagtggt cccaaagatg   3480
gacccccacc cacgaggagc atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc   3540
aagtggattg atgtgacatc tccactgacg taagggatga cgcacaatcc cactatcctt   3600
cgcaagaccc ttcctctata taaggaagtt catttcattt ggagaggaca cgcacgagat   3660
ctgcaggtcg actctagagg atcccctcga ggcgcgccac acgcaaacac aaatacacag   3720
cggccttgcc accatggtgg atctacgcac gctcggctac agccagcagc aacaggagaa   3780
gatcaaaccg aaggttcgtt cgacagtggc gcagcaccac gaggcactgg tcggccacgg   3840
gtttacacac gcgcacatcg ttgcgttaag ccaacacccg gcagcgttag ggaccgtcgc   3900
tgtcaagtat caggacatga tcgcagcgtt gccagaggcg acacacgaag cgatcgttgg   3960
cgtcggcaaa cagtggtccg gcgcacgcgc tctggaggcc ttgctcacgg tggcgggaga   4020
gttgagaggt ccaccgttac agttggacac aggccaactt ctcaagattg caaaacgtgg   4080
cggcgtgacc gcagtggagg cagtgcatgc atggcgcaat gcactgacgg gtgccccgct   4140
caacttgacc ccccagcagg tggtggccat cgccagcaat aatggtggca agcaggcgct   4200
ggagacggtc cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccggagca   4260
ggtggtggcc atcgccagca atattggtgg caagcaggcg ctggagacgg tgcaggcgct   4320
gttgccggtg ctgtgccagg cccacggctt gaccccggag caggtggtgg ccatcgccag   4380
caatattggt ggcaagcagg cgctggagac ggtgcaggcg ctgttgccgg tgctgtgcca   4440
ggcccacggc ttgacccccc agcaggtggt ggccatcgcc agcaataatg gtggcaagca   4500
ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc   4560
ggagcaggtg gtggccatcg ccagccacga tggcggcaag caggcgctgg agacggtcca   4620
gcggctgttg ccggtgctgt gccaggccca cggcttgacc ccccagcagg tggtggccat   4680
cgccagcaat aatggtggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct   4740
gtgccaggcc cacggcttga ccccccagca ggtggtggcc atcgccagca ataatggtgg   4800
caagcaggcg ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt   4860
gacccccag caggtggtgg ccatcgccag caataatggt ggcaagcagg cgctggagac   4920
ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt   4980
ggccatcgcc agcaatattg gtggcaagca ggcgctggag acggtgcagg cgctgttgcc   5040
ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg gtggccatcg ccagcaatat   5100
tggtggcaag caggcgctgg agacggtgca ggcgctgttg ccggtgctgt gccaggccca   5160
cggcttgacc ccccagcagg tggtggccat cgccagcaat aatggtggca agcaggcgct   5220
ggagacggtc cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccccagca   5280
ggtggtggcc atcgccagca ataatggtgg caagcaggcg ctggagacgg tccagcggct   5340
gttgccggtg ctgtgccagg cccacggctt gacccccccag caggtggtgg ccatcgccag   5400
caataatggt ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca   5460
ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc agcaatattg gtggcaagca   5520
ggcgctggag acggtgcagg cgctgttgcc ggtgctgtgc caggcccacg gcttgacccc   5580
ggagcaggtg gtggccatcg ccagccacga tggcggcaag caggcgctgg agacggtcca   5640
gcggctgttg ccggtgctgt gccaggccca cggcttgacc cctcagcagg tggtggccat   5700
cgccagcaat ggcggcggca ggccggcgct ggagagcatt gttgcccagt tatctcgccc   5760
tgatccggga tcctccggat ctaaacgcaa caaaccaggt aaggccactg gcaaaggtaa   5820
accagtgaac aataagtggc tgaacaacgc tggcaaagac ttgggttctc cagtgccgga   5880
tcgcattgcc aacaagctga gagacaaaga attcaagagt tttgatgact ttcgcgcaaa   5940
attctgggaa gaagtgtcta aagatccaga actgtctaaa cagtttagtc gcaacaacaa   6000
tgacagaatg aaagttggta aagctccgaa gacccgcacc caagacgtgt ctggcaaacg   6060
caccagtttt gaattgcatc atgaaaaacc aatctctcag aatggtggcg tttacgacat   6120
ggacaacatc tctgtggtta ctccgaaaag acatatcgat attcatcgcg gtaaaggatc   6180
caccggtgac tgatcggccg actgataagt cgaccaccac caccaccacc actgataaga   6240
gctcgaattt ccccgatcgt tcaaacattt ggcaataaag tttcttaaga ttgaatcctg   6300
ttgccggtct tgcgatgatt atcatataat ttctgttgaa ttacgttaag catgtaataa   6360
```

ttaacatgta atgcatgacg ttatttatga gatgggtttt tatgattaga gtcccgcaat 6420
tatacattta atacgcgata gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc 6480
gcgcggtgtc atctatgtta ctagatcggg aattcgtaat catggtcata gctgtttcct 6540
gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag cataaagtgt 6600
aaagcctggg gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc 6660
gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg 6720
agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg 6780
gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca 6840
gaatcagggg ataacgcagg aaagaac 6867

<210> 290

<211> 220
<212> DNA
<213> Artificial sequence

<220>
<223> TCRB gene region

<400> 290
gccaacagtg tcctaccagc aaggggtcct gtctgccacc atcctctatg agatcctgct 60
agggaaggcc accctgtatg ctgtgctggt cagcgccctt gtgttgatgg ccatggtaag 120
caggagggca ggatggggcc agcaggctgg aggtgacaca ctgacaccaa gcacccagaa 180
gtatagagtc cctgccagga ttggagctgg gcagtaggga 220

<210> 291

<211> 349
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS6857

<400> 291
Met Ala Asn Thr Lys Tyr Asn Glu Glu Phe Leu Leu Tyr Leu Ala Gly
1 5 10 15

Phe Val Asp Gly Asp Gly Ser Ile Val Ala Gln Ile Lys Pro His Gln
20 25 30

Ser Cys Lys Phe Lys His Gln Leu Arg Leu Thr Phe Asn Val Thr Gln
35 40 45

Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
50 55 60

Val Gly His Val Tyr Asp Ser Gly Ser Val Ser Tyr Tyr Gln Leu Ser
65 70 75 80

Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu
85 90 95

Glu Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln
100 105 110

Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr
115 120 125

```
Trp Val Asp Gln Val Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr
    130                 135                 140

Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys
145                 150                 155                 160

Lys Ser Ser Pro Ala Ala Gly Asp Ser Ser Val Ser Asn Ser Glu His
                165                 170                 175

Ile Ala Pro Leu Ser Leu Pro Ser Ser Pro Pro Ser Val Gly Ser Asn
            180                 185                 190

Lys Lys Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Ser Asp Gly Ser
        195                 200                 205

Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Tyr
    210                 215                 220

Leu Ser Leu Thr Phe Cys Val Thr Gln Lys Thr Gln Arg Arg Trp Phe
225                 230                 235                 240

Leu Asp Lys Leu Val Asp Arg Ile Gly Val Gly Tyr Val Gln Asp Arg
                245                 250                 255

Gly Ser Val Ser Asp Tyr Arg Leu Ser Glu Ile Lys Pro Leu Arg Asn
            260                 265                 270

Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala
        275                 280                 285

Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser
    290                 295                 300

Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Val Ala Ala
305                 310                 315                 320

Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala
                325                 330                 335

Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro
            340                 345
```

<210> 292

<211> 4541
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15788

<400> 292

```
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc     60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga    120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg    180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag    240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga    300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc    360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca    420
caaatacaca gcggccttgc caccatggtg gatctacgca cgctcggcta cagccagcag    480
caacaggaga agatcaaacc gaaggttcgt tcgacagtgg cgcagcacca cgaggcactg    540
gtcggccacg ggtttacaca cgcgcacatc gttgcgttaa gccaacaccc ggcagcgtta    600
gggaccgtcg ctgtcaagta tcaggacatg atcgcagcgt tgccagaggc gacacacgaa    660
gcgatcgttg gcgtcggcaa acagtggtcc ggcgcacgcg ctctggaggc cttgctcacg    720
gtggcgggag agttgagagg tccaccgtta cagttggaca caggccaact tctcaagatt    780
gcaaaacgtg gcggcgtgac cgcagtggag gcagtgcatg catggcgcaa tgcactgacg    840
ggtgccccgc tcaacttgac cggagacgcc cgggggatca ggtcacgtgc gtctcggagc    900
attgttgccc agttatctcg ccctgatccg ggatccgaca tcaccaaatc taaaatctct    960
gaaaagatga aaggccaggg cccatccgga aataccaaat ataacgaaga gttcctgctg   1020
tacctggccg gctttgtgga cggtgacggt agcatcgttg ctcagattaa accacatcag   1080
tcttgtaagt ttaaacatca gctacgtttg acctttaacg tgactcaaaa gacccagcgc   1140
cgttggtttc tggacaaact agtggatgaa attggcgttg gtcatgtata cgattctggt   1200
agcgtttcct actaccagtt aagcgaaatc aagccgctgc acaacttcct gactcaactg   1260
cagccgtttc tggaactgaa acagaaacag gcaaacctgg ttctgaaaat tatcgaacag   1320
ctgccgtctg caaaagaatc cccggacaaa ttcctggaag tttgtacctg ggtggatcag   1380
gttgcagctc tgaacgattc taagacgcgt aaaaccactt ctgaaaccgt tcgtgctgtg   1440
ctggacagcc tgagcgagaa gaagaaatcc tccccggcgg ccggtgattc ctctgtttct   1500
aattccgagc acattgctcc tctgtctctg ccttcctctc ctccatctgt tggttctaac   1560
aaaaaattcc tgctgtatct tgctggattt gtggattctg atggctccat cattgctcag   1620
ataaaaccaa atcaatctta caagttcaaa cactacctct ccttgacctt ttgtgtcact   1680
cagaagacac aaagaaggtg gttcttggac aaattggttg atcgtattgg tgtgggctat   1740
gtccaggaca gaggctctgt gtcagactac cgtctgtctg aaattaagcc tcttcgtaac   1800
tttctcaccc aactgcaacc cttcttgaag ctcaaacaga agcaagcaaa tctggttttg   1860
aaaatcatcg agcaactgcc atctgccaag gagtcccctg acaagtttct tgaagtgtgt   1920
acttgggtgg atcaggttgc tgccttgaat gactccaaga ccagaaaaac cacctctgag   1980
actgtgaggg cagttctgga tagcctctct gagaagaaaa agtcctctcc ttagtaactc   2040
gagcctgcag cagctgaagc tttggacttc tggcgcgccc aattcgccct atagtgagtc   2100
gtattacgtc gcgctcactg gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg   2160
ttacccaact taatcgcctt gcagcacatc ccccttcgc cagctggcgt aatagcgaag   2220
aggcccgcac cgaaacgccc ttcccaacag ttgcgcagcc tgaatggcga atgggagcgc   2280
cctgtagcgg cgcattaagc gcggcgggtg tggtggttac gcgcagcgtg accgctacac   2340
ttgccagcgc cctagcgccc gctcctttcg ctttcttccc ttcctttctc gccacgttcg   2400
ccggctttcc ccgtcaagct ctaaatcggg ggctcccttt agggttccga tttagtgctt   2460
tacggcacct cgaccccaaa aaacttgatt agggtgatgg ttggcctgta gtgggccata   2520
gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta atagtggact   2580
cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg   2640
gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc   2700
gaattttaac aaaatattaa cgcttacaat ttaggtggca cttttcgggg aaatgtgcgc   2760
ggaaccccta tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa   2820
taaccctgat aaatgcttca ataatattga aaaaggaaga gtatgagtat tcaacatttc   2880
cgtgtcgccc ttattccctt ttttgcggca ttttgccttc ctgtttttgc tcacccagaa   2940
acgctggtga aagtaaaaga tgctgaagat cagttgggtg cacgagtggg ttacatcgaa   3000
ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg   3060
atgagcactt ttaaagttct gctatgtggc gcggtattat cccgtattga cgccgggcaa   3120
gagcaactcg gtcgccgcat acactattct cagaatgact tggttgagta ctcaccagtc   3180
acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc   3240
atgagtgata acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta   3300
accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag   3360
ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgcctgtagc aatggcaaca   3420
acgttgcgca aactattaac tggcgaacta cttactctag cttcccggca acaattaata   3480
gactggatgg aggcggataa agttgcagga ccacttctgc gctcggccct tccggctggc   3540
tggtttattg ctgataaatc tggagccggt gagcgtggtt ctcgcggtat cattgcagca   3600
ctggggccag atggtaagcc ctcccgtatc gtagttatct acacgacggg gagtcaggca   3660
actatggatg aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg   3720
taactgtcag accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa   3780
tttaaaagga tctaggtgaa gatccttttt gataatctca tgaccaaaat cccttaacgt   3840
gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat   3900
ccttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg   3960
gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga   4020
gcgcagatac caaatactgt tcttctagtg tagccgtagt taggccacca cttcaagaac   4080
```

```
tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt   4140
ggcgataagt cgtgtcttac cgggttggac tcaagacgat agttaccgga taaggcgcag   4200
cggtcgggct gaacgggggg ttcgtgcaca gcccagct tggagcgaac gacctacacc   4260
gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaag   4320
gcggacaggt atccggtaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca   4380
gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt   4440
cgatttttgt gatgctcgtc agggggggcgg agcctatgga aaaacgccag caacgcggcc   4500
tttttacggt tcctggcctt ttgctggcct tttgctcaca t                       4541
```

<210> 293

<211> 349
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS12083

<400> 293

```
Met Ala Asn Thr Lys Tyr Asn Glu Glu Phe Leu Leu Tyr Leu Ala Gly
1               5                  10                  15

Phe Val Asp Gly Asp Gly Ser Ile Val Ala Gln Ile Lys Pro His Gln
            20                  25                  30

Ser Cys Lys Phe Lys His Gln Leu Arg Leu Thr Phe Asn Val Thr Gln
        35                  40                  45

Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
    50                  55                  60

Val Gly His Val Tyr Asp Ser Gly Ser Val Ser Tyr Tyr Gln Leu Ser
65                  70                  75                  80

Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu
                85                  90                  95

Glu Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln
            100                 105                 110

Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr
        115                 120                 125

Trp Val Asp Gln Val Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr
    130                 135                 140

Thr Ser Glu Thr Cys Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys
145                 150                 155                 160

Lys Ser Ser Pro Ala Ala Gly Asp Ser Ser Val Ser Asn Ser Glu His
                165                 170                 175

Ile Ala Pro Leu Ser Leu Pro Ser Ser Pro Pro Ser Val Gly Ser Asn
            180                 185                 190
```

Lys Lys Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Ser Asp Gly Ser
195 200 205

Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Tyr
210 215 220

Leu Ser Leu Thr Phe Cys Val Thr Gln Lys Thr Gln Arg Arg Trp Phe
225 230 235 240

Leu Asp Lys Leu Val Asp Arg Ile Gly Val Gly Tyr Val Gln Asp Arg
245 250 255

Gly Ser Val Ser Asp Tyr Arg Leu Ser Glu Ile Lys Pro Leu Arg Asn
260 265 270

Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala
275 280 285

Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser
290 295 300

Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Val Ala Ala
305 310 315 320

Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala
325 330 335

Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro
340 345

<210> 294

<211> 349
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS12195

<400> 294
Met Ala Asn Thr Lys Tyr Asn Glu Glu Phe Leu Leu Tyr Leu Ala Gly
1 5 10 15

Phe Val Asp Gly Asp Gly Ser Ile Val Ala Gln Ile Lys Pro His Gln
20 25 30

Ser Cys Lys Phe Lys His Gln Leu Arg Leu Thr Phe Asn Val Thr Gln
35 40 45

Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
50 55 60

Val Gly His Val Tyr Asp Ser Gly Ser Val Ser Tyr Tyr Gln Leu Ser

65 70 75 80

Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu
85 90 95

Glu Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln
100 105 110

Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr
115 120 125

Trp Val Asp Gln Val Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr
130 135 140

Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys
145 150 155 160

Lys Ser Ser Pro Ala Ala Gly Asp Ser Ser Val Ser Asn Ser Glu His
165 170 175

Ile Ala Pro Leu Ser Leu Pro Ser Ser Pro Pro Ser Val Gly Ser Asn
180 185 190

Lys Lys Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Ser Asp Gly Ser
195 200 205

Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Tyr
210 215 220

Leu Ser Leu Thr Phe Cys Val Thr Gln Lys Thr Gln Arg Arg Trp Phe
225 230 235 240

Leu Asp Lys Leu Val Asp Arg Ile Gly Val Gly Tyr Val Gln Asp Arg
245 250 255

Gly Ser Val Ser Asp Tyr Arg Leu Ser Glu Ile Lys Pro Leu Arg Asn
260 265 270

Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala
275 280 285

Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser
290 295 300

Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Val Ala Ala
305 310 315 320

Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Cys Arg Ala
325 330 335

Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro

340 345

<210> 295

<211> 4541
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15789

<400> 295
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggtg gatctacgca cgctcggcta cagccagcag 480
caacaggaga agatcaaacc gaaggttcgt tcgacagtgg cgcagcacca cgaggcactg 540
gtcggccacg ggtttacaca cgcgcacatc gttgcgttaa gccaacaccc ggcagcgtta 600
gggaccgtcg ctgtcaagta tcaggacatg atcgcagcgt tgccagaggc gacacacgaa 660
gcgatcgttg gcgtcggcaa acagtggtcc ggcgcacgcg ctctggaggc cttgctcacg 720
gtggcgggag agttgagagg tccaccgtta cagttggaca caggccaact tctcaagatt 780
gcaaaacgtg gcggcgtgac cgcagtggag gcagtgcatg catggcgcaa tgcactgacg 840
ggtgccccgc tcaacttgac cggagacgcc cgggggatca ggtcacgtgc gtctcggagc 900
attgttgccc agttatctcg ccctgatccg ggatccgaca tcaccaaatc taaaatctct 960
gaaaagatga aaggccaggg cccatccgga aataccaaat ataacgaaga gttcctgctg 1020
tacctggccg gctttgtgga cggtgacggt agcatcgttg ctcagattaa accacatcag 1080
tcttgtaagt ttaaacatca gctacgtttg acctttaacg tgactcaaaa gacccagcgc 1140
cgttggtttc tggacaaact agtggatgaa attggcgttg gtcatgtata cgattctggt 1200
agcgtttcct actaccagtt aagcgaaatc aagccgctgc acaacttcct gactcaactg 1260
cagccgtttc tggaactgaa acagaaacag gcaaacctgg ttctgaaaat tatcgaacag 1320
ctgccgtctg caaaagaatc cccggacaaa ttcctggaag tttgtacctg ggtggatcag 1380
gttgcagctc tgaacgattc taagacgcgt aaaaccactt ctgaaacctg tcgtgctgtg 1440
ctggacagcc tgagcgagaa gaagaaatcc tccccggcgg ccggtgattc ctctgtttct 1500
aattccgagc acattgctcc tctgtctctg ccttcctctc ctccatctgt tggttctaac 1560
aaaaaattcc tgctgtatct tgctggattt gtggattctg atggctccat cattgctcag 1620
ataaaaccaa atcaatctta caagttcaaa cactacctct ccttgacctt ttgtgtcact 1680
cagaagacac aaagaaggtg gttcttggac aaattggttg atcgtattgg tgtgggctat 1740
gtccaggaca gaggctctgt gtcagactac cgtctgtctg aaattaagcc tcttcgtaac 1800
tttctcaccc aactgcaacc cttcttgaag ctcaaacaga agcaagcaaa tctggttttg 1860
aaaatcatcg agcaactgcc atctgccaag gagtcccctg acaagtttct tgaagtgtgt 1920
acttgggtgg atcaggttgc tgccttgaat gactccaaga ccagaaaaac cacctctgag 1980
actgtgaggg cagttctgga tagcctctct gagaagaaaa agtcctctcc ttagtaactc 2040
gagcctgcag cagctgaagc tttggacttc tggcgcgccc aattcgccct atagtgagtc 2100
gtattacgtc gcgctcactg gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg 2160
ttacccaact taatcgcctt gcagcacatc ccccttcgc cagctggcgt aatagcgaag 2220
aggcccgcac cgaaacgccc ttcccaacag ttgcgcagcc tgaatggcga atgggagcgc 2280
cctgtagcgg cgcattaagc gcggcgggtg tggtggttac gcgcagcgtg accgctacac 2340
ttgccagcgc cctagcgccc gctcctttcg ctttcttccc ttcctttctc gccacgttcg 2400
ccggctttcc ccgtcaagct ctaaatcggg ggctcccttt agggttccga tttagtgctt 2460
tacggcacct cgaccccaaa aaacttgatt agggtgatgg ttggcctgta gtgggccata 2520
gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta atagtggact 2580
cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg 2640
gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc 2700
gaattttaac aaaatattaa cgcttacaat ttaggtggca cttttcgggg aaatgtgcgc 2760
ggaacccctа tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa 2820
taaccctgat aaatgcttca ataatattga aaaaggaaga gtatgagtat tcaacatttc 2880
cgtgtcgccc ttattccctt ttttgcggca ttttgccttc ctgtttttgc tcacccagaa 2940
acgctggtga aagtaaaaga tgctgaagat cagttgggtg cacgagtggg ttacatcgaa 3000
ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg 3060
atgagcactt ttaaagttct gctatgtggc gcggtattat cccgtattga cgccgggcaa 3120
gagcaactcg gtcgccgcat acactattct cagaatgact tggttgagta ctcaccagtc 3180
acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc 3240
atgagtgata acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta 3300
accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag 3360

ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgcctgtagc aatggcaaca 3420
acgttgcgca aactattaac tggcgaacta cttactctag cttcccggca acaattaata 3480
gactggatgg aggcggataa agttgcagga ccacttctgc gctcggccct tccggctggc 3540
tggtttattg ctgataaatc tggagccggt gagcgtggtt ctcgcggtat cattgcagca 3600
ctggggccag atggtaagcc ctcccgtatc gtagttatct acacgacggg gagtcaggca 3660
actatggatg aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg 3720
taactgtcag accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa 3780
tttaaaagga tctaggtgaa gatccttttt gataatctca tgaccaaaat cccttaacgt 3840
gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat 3900
ccttttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg 3960
gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga 4020
gcgcagatac caaatactgt tcttctagtg tagccgtagt taggccacca cttcaagaac 4080
tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt 4140
ggcgataagt cgtgtcttac cgggttggac tcaagacgat agttaccgga taaggcgcag 4200
cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc 4260
gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaag 4320
gcggacaggt atccggtaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca 4380
gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt 4440
cgatttttgt gatgctcgtc agggggggcgg agcctatgga aaaacgccag caacgcggcc 4500
tttttacggt tcctggcctt ttgctggcct tttgctcaca t 4541

<210> 296

<211> 4541
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15790

<400> 296
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggtg gatctacgca cgctcggcta cagccagcag 480
caacaggaga agatcaaacc gaaggttcgt tcgacagtgg cgcagcacca cgaggcactg 540
gtcggccacg ggtttacaca cgcgcacatc gttgcgttaa gccaacaccc ggcagcgtta 600
gggaccgtcg ctgtcaagta tcaggacatg atcgcagcgt tgccagaggc gacacacgaa 660
gcgatcgttg gcgtcggcaa acagtggtcc ggcgcacgcg ctctggaggc cttgctcacg 720
gtggcgggag agttgagagg tccaccgtta cagttggaca caggccaact tctcaagatt 780
gcaaaacgtg gcggcgtgac cgcagtggag gcagtgcatg catggcgcaa tgcactgacg 840
ggtgccccgc tcaacttgac cggagacgcc cgggggatca ggtcacgtgc gtctcggagc 900
attgttgccc agttatctcg ccctgatccg ggatccgaca tcaccaaatc taaaatctct 960
gaaaagatga aaggccaggg cccatccgga aataccaaat ataacgaaga gttcctgctg 1020
tacctggccg gctttgtgga cggtgacggt agcatcgttg ctcagattaa accacatcag 1080
tcttgtaagt ttaaacatca gctacgtttg acctttaacg tgactcaaaa gacccagcgc 1140
cgttggtttc tggacaaact agtggatgaa attggcgttg gtcatgtata cgattctggt 1200
agcgtttcct actaccagtt aagcgaaatc aagccgctgc acaacttcct gactcaactg 1260
cagccgtttc tggaactgaa acagaaacag gcaaacctgg ttctgaaaat tatcgaacag 1320
ctgccgtctg caaaagaatc cccggacaaa ttcctggaag tttgtacctg ggtggatcag 1380
gttgcagctc tgaacgattc taagacgcgt aaaaccactt ctgaaaccgt tcgtgctgtg 1440
ctggacagcc tgagcgagaa gaagaaatcc tccccggcgg ccggtgattc ctctgtttct 1500
aattccgagc acattgctcc tctgtctctg ccttcctctc ctccatctgt tggttctaac 1560
aaaaaattcc tgctgtatct tgctggattt gtggattctg atggctccat cattgctcag 1620
ataaaaccaa atcaatctta caagttcaaa cactacctct ccttgacctt ttgtgtcact 1680
cagaagacac aaagaaggtg gttcttggac aaattggttg atcgtattgg tgtgggctat 1740
gtccaggaca gaggctctgt gtcagactac cgtctgtctg aaattaagcc tcttcgtaac 1800
tttctcaccc aactgcaacc cttcttgaag ctcaaacaga agcaagcaaa tctggttttg 1860
aaaatcatcg agcaactgcc atctgccaag gagtcccctg acaagtttct tgaagtgtgt 1920
acttgggtgg atcaggttgc tgccttgaat gactccaaga ccagaaaaac cacctctgag 1980
acttgtaggg cagttctgga tagcctctct gagaagaaaa agtcctctcc ttagtaactc 2040
gagcctgcag cagctgaagc tttggacttc tggcgcgccc aattcgccct atagtgagtc 2100
gtattacgtc gcgctcactg gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg 2160

ttacccaact taatcgcctt gcagcacatc ccccttcgc cagctggcgt aatagcgaag 2220
aggcccgcac cgaaacgccc ttcccaacag ttgcgcagcc tgaatggcga atgggagcgc 2280
cctgtagcgg cgcattaagc gcggcgggtg tggtggttac gcgcagcgtg accgctacac 2340
ttgccagcgc cctagcgccc gctcctttcg ctttcttccc ttcctttctc gccacgttcg 2400
ccggctttcc ccgtcaagct ctaaatcggg ggctcccttt agggttccga tttagtgctt 2460
tacggcacct cgaccccaaa aaacttgatt agggtgatgg ttggcctgta gtgggccata 2520
gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta atagtggact 2580
cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg 2640
gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc 2700
gaattttaac aaaatattaa cgcttacaat ttaggtggca cttttcgggg aaatgtgcgc 2760
ggaaccccta tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa 2820
taaccctgat aaatgcttca ataatattga aaaaggaaga gtatgagtat tcaacatttc 2880
cgtgtcgccc ttattcccct ttttgcggca ttttgccttc ctgtttttgc tcacccagaa 2940
acgctggtga aagtaaaaga tgctgaagat cagttgggtg cacgagtggg ttacatcgaa 3000
ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg 3060
atgagcactt ttaaagttct gctatgtggc gcggtattat cccgtattga cgccgggcaa 3120
gagcaactcg gtcgccgcat acactattct cagaatgact tggttgagta ctcaccagtc 3180
acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc 3240
atgagtgata acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta 3300
accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag 3360
ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgcctgtagc aatggcaaca 3420
acgttgcgca aactattaac tggcgaacta cttactctag cttcccggca acaattaata 3480
gactggatgg aggcggataa agttgcagga ccacttctgc gctcggccct tccggctggc 3540
tggtttattg ctgataaatc tggagccggt gagcgtggtt ctcgcggtat cattgcagca 3600
ctggggccag atggtaagcc ctcccgtatc gtagttatct acacgacggg gagtcaggca 3660
actatggatg aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg 3720
taactgtcag accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa 3780
tttaaaagga tctaggtgaa gatccttttt gataatctca tgaccaaaat cccttaacgt 3840
gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat 3900
ccttttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg 3960
gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga 4020
gcgcagatac caaatactgt tcttctagtg tagccgtagt taggccacca cttcaagaac 4080
tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt 4140
ggcgataagt cgtgtcttac cgggttggac tcaagacgat agttaccgga taaggcgcag 4200
cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc 4260
gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaag 4320
gcggacaggt atccggtaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca 4380
gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt 4440
cgatttttgt gatgctcgtc agggggggcgg agcctatgga aaaacgccag caacgcggcc 4500
tttttacggt tcctggcctt ttgctggcct tttgctcaca t 4541

<210> 297

<211> 1410
<212> DNA
<213> Artificial sequence

<220>
<223> TCRB02A1 RVD

<400> 297
gcagcctcaa cttgaccccc cagcaggtgg tggccatcgc cagcaataat ggtggcaagc 60
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 120
cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg gagacggtcc 180
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag gtggtggcca 240
tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 300
tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc aataatggtg 360
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 420
tgacccccca gcaggtggtg gccatcgcca gcaataatgg tggcaagcag gcgctggaga 480
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgacccccc cagcaggtgg 540
tggccatcgc cagcaatggc ggtggcaagc aggcgctgga gacggtccag cggctgttgc 600
cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc gccagccacg 660
atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg tgccaggccc 720
acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc aagcaggcgc 780
tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg acccccccagc 840
aggtggtggc catcgccagc aataatggtg gcaagcaggc gctggagacg gtccagcggc 900
tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg gccatcgcca 960
gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc 1020

aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaataat ggtggcaagc 1080
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 1140
cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg gagacggtcc 1200
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag gtggtggcca 1260
tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 1320
tgtgccaggc ccacggcttg acccctcagc aggtggtggc catcgccagc aatggcggcg 1380
gcaggccggc gctggagagc attgtgatgc 1410

<210> 298

<211> 1410
<212> DNA
<213> Artificial sequence

<220>
<223> TCRB02A2 RVD

<400> 298
gcagcctcaa cttgaccccc cagcaggtgg tggccatcgc cagcaataat ggtggcaagc 60
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 120
cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg gagacggtcc 180
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag gtggtggcca 240
tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 300
tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc aataatggtg 360
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 420
tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag gcgctggaga 480
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg 540
tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag cggctgttgc 600
cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc gccagccacg 660
atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg tgccaggccc 720
acggcttgac cccccagcag gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc 780
tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg acccccccagc 840
aggtggtggc catcgccagc aataatggtg gcaagcaggc gctggagacg gtccagcggc 900
tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg gccatcgcca 960
gcaataatgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc 1020
aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc ggtggcaagc 1080
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 1140
cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg gagacggtcc 1200
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag gtggtggcca 1260
tcgccagcaa tattggtggc aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc 1320
tgtgccaggc ccacggcttg acccctcagc aggtggtggc catcgccagc aatggcggcg 1380
gcaggccggc gctggagagc attgtgatgc 1410

<210> 299

<211> 1410
<212> DNA
<213> Artificial sequence

<220>
<223> TCRB02A3 RVD

<400> 299
gcagcctcaa cttgaccccc cagcaggtgg tggccatcgc cagcaataat ggtggcaagc 60
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 120
cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc 180
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag gtggtggcca 240
tcgccagcaa tattggtggc aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc 300
tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc aatggcggtg 360
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 420
tgacccccca gcaggtggtg gccatcgcca gcaataatgg tggcaagcag gcgctggaga 480
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg 540
tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag cggctgttgc 600
cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc gccagcaatg 660
gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg tgccaggccc 720
acggcttgac cccccagcag gtggtggcca tcgccagcaa taatggtggc aagcaggcgc 780

tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg acccccccagc 840
aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg gtccagcggc 900
tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg gccatcgcca 960
gcaataatgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc 1020
aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat ggcggcaagc 1080
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 1140
cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc 1200
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag gtggtggcca 1260
tcgccagcaa taatggtggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 1320
tgtgccaggc ccacggcttg acccctcagc aggtggtggc catcgccagc aatggcggcg 1380
gcaggccggc gctggagagc attgtgatgc 1410

<210> 300

<211> 5885
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15791

<400> 300
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggtg gatctacgca cgctcggcta cagccagcag 480
caacaggaga agatcaaacc gaaggttcgt tcgacagtgg cgcagcacca cgaggcactg 540
gtcggccacg ggtttacaca cgcgcacatc gttgcgttaa gccaacaccc ggcagcgtta 600
gggaccgtcg ctgtcaagta tcaggacatg atcgcagcgt tgccagaggc gacacacgaa 660
gcgatcgttg gcgtcggcaa acagtggtcc ggcgcacgcg ctctggaggc cttgctcacg 720
gtggcgggag agttgagagg tccaccgtta cagttggaca caggccaact tctcaagatt 780
gcaaaacgtg gcggcgtgac cgcagtggag gcagtgcatg catggcgcaa tgcactgacg 840
ggtgccccgc tcaacttgac cccccagcag gtggtggcca tcgccagcaa taatggtggc 900
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 960
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 1020
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg 1080
gccatcgcca gcaatggcgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg 1140
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaataat 1200
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 1260
ggcttgaccc cccagcaggt ggtggccatc gccagcaata atggtggcaa gcaggcgctg 1320
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 1380
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 1440
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 1500
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 1560
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag 1620
gcgctggaga cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc 1680
cagcaggtgg tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag 1740
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 1800
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 1860
tgccaggccc acggcttgac cccccagcag gtggtggcca tcgccagcaa taatggtggc 1920
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 1980
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 2040
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 2100
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 2160
gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg tggccatcgc cagcaatggc 2220
ggcggcaggc cggcgctgga gagcattgtt gcccagttat ctcgccctga tccgggatcc 2280
gacatcacca aatctaaaat ctctgaaaag atgaaaggcc agggcccatc cggaaatacc 2340
aaatataacg aagagttcct gctgtacctg gccggctttg tggacggtga cggtagcatc 2400
gttgctcaga ttaaaccaca tcagtcttgt aagtttaaac atcagctacg tttgaccttt 2460
aacgtgactc aaaagaccca gcgccgttgg tttctggaca aactagtgga tgaaattggc 2520
gttggtcatg tatacgattc tggtagcgtt tcctactacc agttaagcga aatcaagccg 2580
ctgcacaact tcctgactca actgcagccg tttctggaac tgaaacagaa acaggcaaac 2640
ctggttctga aaattatcga acagctgccg tctgcaaaag aatccccgga caaattcctg 2700

gaagtttgta cctgggtgga tcaggttgca gctctgaacg attctaagac gcgtaaaacc 2760
acttctgaaa ccgttcgtgc tgtgctggac agcctgagcg agaagaagaa atcctccccg 2820
gcggccggtg attcctctgt ttctaattcc gagcacattg ctcctctgtc tctgccttcc 2880
tctcctccat ctgttggttc taacaaaaaa ttcctgctgt atcttgctgg atttgtggat 2940
tctgatggct ccatcattgc tcagataaaa ccaaatcaat cttacaagtt caaacactac 3000
ctctccttga ccttttgtgt cactcagaag acacaaagaa ggtggttctt ggacaaattg 3060
gttgatcgta ttggtgtggg ctatgtccag gacagaggct ctgtgtcaga ctaccgtctg 3120
tctgaaatta agcctcttcg taactttctc acccaactgc aacccttctt gaagctcaaa 3180
cagaagcaag caaatctggt tttgaaaatc atcgagcaac tgccatctgc caaggagtcc 3240
cctgacaagt ttcttgaagt gtgtacttgg gtggatcagg ttgctgcctt gaatgactcc 3300
aagaccagaa aaaccacctc tgagactgtg agggcagttc tggatagcct ctctgagaag 3360
aaaaagtcct ctccttagta actcgagcct gcagcagctg aagctttgga cttctggcgc 3420
gcccaattcg ccctatagtg agtcgtatta cgtcgcgctc actggccgtc gttttacaac 3480
gtcgtgactg ggaaaaccct ggcgttaccc aacttaatcg ccttgcagca catccccctt 3540
tcgccagctg gcgtaatagc gaagaggccc gcaccgaaac gcccttccca acagttgcgc 3600
agcctgaatg gcgaatggga gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg 3660
ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct 3720
tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc 3780
ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg 3840
atggttggcc tgtagtgggc catagccctg atagacggtt tttcgccctt tgacgttgga 3900
gtccacgttc tttaatagtg gactcttgtt ccaaactgga acaacactca accctatctc 3960
ggtctattct tttgatttat aagggatttt gccgatttcg gcctattggt taaaaaatga 4020
gctgatttaa caaaaattta acgcgaattt taacaaaata ttaacgctta caatttaggt 4080
ggcacttttc ggggaaatgt gcgcggaacc cctatttgtt tatttttcta aatacattca 4140
aatatgtatc cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg 4200
aagagtatga gtattcaaca tttccgtgtc gcccttattc cctttttgc ggcattttgc 4260
cttcctgttt ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg 4320
ggtgcacgag tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt 4380
cgccccgaag aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta 4440
ttatcccgta ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat 4500
gacttggttg agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga 4560
gaattatgca gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca 4620
acgatcggag gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact 4680
cgccttgatc gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc 4740
acgatgcctg tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact 4800
ctagcttccc ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt 4860
ctgcgctcgg cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt 4920
ggttctcgcg gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt 4980
atctacacga cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata 5040
ggtgcctcac tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag 5100
attgatttaa aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat 5160
ctcatgacca aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa 5220
aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca 5280
aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt 5340
ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgttcttct agtgtagccg 5400
tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatc 5460
ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga 5520
cgatagttac cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc 5580
agcttggagc gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc 5640
gccacgcttc ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca 5700
ggagagcgca cgagggagct tccaggggga aacgcctggt atctttatag tcctgtcggg 5760
tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta 5820
tggaaaaacg ccagcaacgc ggccttttta cggttcctgg ccttttgctg gccttttgct 5880
cacat 5885

<210> 301

<211> 5885
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15792

<400> 301
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120

gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggtg gatctacgca cgctcggcta cagccagcag 480
caacaggaga agatcaaacc gaaggttcgt tcgacagtgg cgcagcacca cgaggcactg 540
gtcggccacg ggtttacaca cgcgcacatc gttgcgttaa gccaacaccc ggcagcgtta 600
gggaccgtcg ctgtcaagta tcaggacatg atcgcagcgt tgccagaggc gacacacgaa 660
gcgatcgttg gcgtcggcaa acagtggtcc ggcgcacgcg ctctggaggc cttgctcacg 720
gtggcgggag agttgagagg tccaccgtta cagttggaca caggccaact tctcaagatt 780
gcaaaacgtg gcggcgtgac cgcagtggag gcagtgcatg catggcgcaa tgcactgacg 840
ggtgccccgc tcaacttgac cccccagcag gtggtggcca tcgccagcaa taatggtggc 900
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 960
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 1020
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg 1080
gccatcgcca gcaatggcgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg 1140
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaataat 1200
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 1260
ggcttgaccc cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg 1320
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 1380
gtggtggcca tcgccagcaa taatggtggc aagcaggcgc tggagacggt ccagcggctg 1440
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 1500
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 1560
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 1620
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc 1680
cagcaggtgg tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag 1740
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 1800
gccagcaata atggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 1860
tgccaggccc acggcttgac cccccagcag gtggtggcca tcgccagcaa tggcggtggc 1920
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 1980
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 2040
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 2100
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 2160
gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg tggccatcgc cagcaatggc 2220
ggcggcaggc cggcgctgga gagcattgtt gcccagttat ctcgccctga tccgggatcc 2280
gacatcacca aatctaaaat ctctgaaaag atgaaaggcc agggcccatc cggaaatacc 2340
aaatataacg aagagttcct gctgtacctg gccggctttg tggacggtga cggtagcatc 2400
gttgctcaga ttaaaccaca tcagtcttgt aagtttaaac atcagctacg tttgaccttt 2460
aacgtgactc aaaagaccca gcgccgttgg tttctggaca aactagtgga tgaaattggc 2520
gttggtcatg tatacgattc tggtagcgtt tcctactacc agttaagcga aatcaagccg 2580
ctgcacaact tcctgactca actgcagccg tttctggaac tgaaacagaa acaggcaaac 2640
ctggttctga aaattatcga acagctgccg tctgcaaaag aatccccgga caaattcctg 2700
gaagtttgta cctgggtgga tcaggttgca gctctgaacg attctaagac gcgtaaaacc 2760
acttctgaaa ccgttcgtgc tgtgctggac agcctgagcg agaagaagaa atcctccccg 2820
gcggccggtg attcctctgt ttctaattcc gagcacattg ctcctctgtc tctgccttcc 2880
tctcctccat ctgttggttc taacaaaaaa ttcctgctgt atcttgctgg atttgtggat 2940
tctgatggct ccatcattgc tcagataaaa ccaaatcaat cttacaagtt caaacactac 3000
ctctccttga ccttttgtgt cactcagaag acacaaagaa ggtggttctt ggacaaattg 3060
gttgatcgta ttggtgtggg ctatgtccag gacagaggct ctgtgtcaga ctaccgtctg 3120
tctgaaatta agcctcttcg taactttctc acccaactgc aacccttctt gaagctcaaa 3180
cagaagcaag caaatctggt tttgaaaatc atcgagcaac tgccatctgc caaggagtcc 3240
cctgacaagt ttcttgaagt gtgtacttgg gtggatcagg ttgctgcctt gaatgactcc 3300
aagaccagaa aaaccacctc tgagactgtg agggcagttc tggatagcct ctctgagaag 3360
aaaaagtcct ctccttagta actcgagcct gcagcagctg aagctttgga cttctggcgc 3420
gcccaattcg ccctatagtg agtcgtatta cgtcgcgctc actggccgtc gttttacaac 3480
gtcgtgactg ggaaaaccct ggcgttaccc aacttaatcg ccttgcagca catccccctt 3540
tcgccagctg gcgtaatagc gaagaggccc gcaccgaaac gcccttccca acagttgcgc 3600
agcctgaatg gcgaatggga gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg 3660
ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct 3720
tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc 3780
ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg 3840
atggttggcc tgtagtgggc catagccctg atagacggtt tttcgccctt tgacgttgga 3900
gtccacgttc tttaatagtg gactcttgtt ccaaactgga acaacactca accctatctc 3960
ggtctattct tttgatttat aagggatttt gccgatttcg gcctattggt taaaaaatga 4020
gctgatttaa caaaaattta acgcgaattt taacaaaata ttaacgctta caatttaggt 4080
ggcacttttc ggggaaatgt gcgcggaacc cctatttgtt tatttttcta aatacattca 4140
aatatgtatc cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg 4200

aagagtatga gtattcaaca tttccgtgtc gcccttattc ccttttttgc ggcattttgc 4260
cttcctgttt ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg 4320
ggtgcacgag tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt 4380
cgccccgaag aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta 4440
ttatcccgta ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat 4500
gacttggttg agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga 4560
gaattatgca gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca 4620
acgatcggag gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact 4680
cgccttgatc gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc 4740
acgatgcctg tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact 4800
ctagcttccc ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt 4860
ctgcgctcgg cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt 4920
ggttctcgcg gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt 4980
atctacacga cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata 5040
ggtgcctcac tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag 5100
attgatttaa aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat 5160
ctcatgacca aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa 5220
aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca 5280
aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt 5340
ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgttcttct agtgtagccg 5400
tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatc 5460
ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga 5520
cgatagttac cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc 5580
agcttggagc gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc 5640
gccacgcttc ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca 5700
ggagagcgca cgagggagct tccaggggga aacgcctggt atctttatag tcctgtcggg 5760
tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta 5820
tggaaaaacg ccagcaacgc ggccttttta cggttcctgg ccttttgctg gccttttgct 5880
cacat 5885

<210> 302

<211> 5885
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15793

<400> 302
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggtg gatctacgca cgctcggcta cagccagcag 480
caacaggaga agatcaaacc gaaggttcgt tcgacagtgg cgcagcacca cgaggcactg 540
gtcggccacg ggtttacaca cgcgcacatc gttgcgttaa gccaacaccc ggcagcgtta 600
gggaccgtcg ctgtcaagta tcaggacatg atcgcagcgt tgccagaggc gacacacgaa 660
gcgatcgttg gcgtcggcaa acagtggtcc ggcgcacgcg ctctggaggc cttgctcacg 720
gtggcgggag agttgagagg tccaccgtta cagttggaca caggccaact tctcaagatt 780
gcaaaacgtg gcggcgtgac cgcagtggag gcagtgcatg catggcgcaa tgcactgacg 840
ggtgccccgc tcaacttgac cccccagcag gtggtggcca tcgccagcaa taatggtggc 900
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 960
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 1020
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 1080
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 1140
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 1200
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 1260
ggcttgaccc cccagcaggt ggtggccatc gccagcaata atggtggcaa gcaggcgctg 1320
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag 1380
gtggtggcca tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg 1440
ttgccggtgc tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc 1500
aatggcggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 1560
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaataatgg tggcaagcag 1620

gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc 1680
cagcaggtgg tggccatcgc cagcaatggc ggtggcaagc aggcgctgga gacggtccag 1740
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 1800
gccagcaata atggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 1860
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 1920
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 1980
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 2040
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg 2100
gccatcgcca gcaataatgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg 2160
gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg tggccatcgc cagcaatggc 2220
ggcggcaggc cggcgctgga gagcattgtt gcccagttat ctcgccctga tccgggatcc 2280
gacatcacca aatctaaaat ctctgaaaag atgaaaggcc agggcccatc cggaaatacc 2340
aaatataacg aagagttcct gctgtacctg gccggctttg tggacggtga cggtagcatc 2400
gttgctcaga ttaaaccaca tcagtcttgt aagtttaaac atcagctacg tttgaccttt 2460
aacgtgactc aaaagaccca gcgccgttgg tttctggaca aactagtgga tgaaattggc 2520
gttggtcatg tatacgattc tggtagcgtt tcctactacc agttaagcga aatcaagccg 2580
ctgcacaact tcctgactca actgcagccg tttctggaac tgaaacagaa acaggcaaac 2640
ctggttctga aaattatcga acagctgccg tctgcaaaag aatccccgga caaattcctg 2700
gaagtttgta cctgggtgga tcaggttgca gctctgaacg attctaagac gcgtaaaacc 2760
acttctgaaa ccgttcgtgc tgtgctggac agcctgagcg agaagaagaa atcctccccg 2820
gcggccggtg attcctctgt ttctaattcc gagcacattg ctcctctgtc tctgccttcc 2880
tctcctccat ctgttggttc taacaaaaaa ttcctgctgt atcttgctgg atttgtggat 2940
tctgatggct ccatcattgc tcagataaaa ccaaatcaat cttacaagtt caaacactac 3000
ctctccttga ccttttgtgt cactcagaag acacaaagaa ggtggttctt ggacaaattg 3060
gttgatcgta ttggtgtggg ctatgtccag gacagaggct ctgtgtcaga ctaccgtctg 3120
tctgaaatta agcctcttcg taactttctc acccaactgc aacccttctt gaagctcaaa 3180
cagaagcaag caaatctggt tttgaaaatc atcgagcaac tgccatctgc caaggagtcc 3240
cctgacaagt ttcttgaagt gtgtacttgg gtggatcagg ttgctgcctt gaatgactcc 3300
aagaccagaa aaaccacctc tgagactgtg agggcagttc tggatagcct ctctgagaag 3360
aaaaagtcct ctccttagta actcgagcct gcagcagctg aagctttgga cttctggcgc 3420
gcccaattcg ccctatagtg agtcgtatta cgtcgcgctc actggccgtc gttttacaac 3480
gtcgtgactg ggaaaaccct ggcgttaccc aacttaatcg ccttgcagca catccccctt 3540
tcgccagctg gcgtaatagc gaagaggccc gcaccgaaac gcccttccca acagttgcgc 3600
agcctgaatg gcgaatggga gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg 3660
ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct 3720
tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc 3780
ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg 3840
atggttggcc tgtagtgggc catagccctg atagacggtt tttcgccctt tgacgttgga 3900
gtccacgttc tttaatagtg gactcttgtt ccaaactgga acaacactca accctatctc 3960
ggtctattct tttgatttat aagggatttt gccgatttcg gcctattggt taaaaaatga 4020
gctgatttaa caaaaattta acgcgaattt taacaaaata ttaacgctta caatttaggt 4080
ggcacttttc ggggaaatgt gcgcggaacc cctatttgtt tatttttcta aatacattca 4140
aatatgtatc cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg 4200
aagagtatga gtattcaaca tttccgtgtc gcccttattc cctttttgc ggcattttgc 4260
cttcctgttt ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg 4320
ggtgcacgag tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt 4380
cgccccgaag aacgttttcc aatgatgagc actttaaag ttctgctatg tggcgcggta 4440
ttatcccgta ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat 4500
gacttggttg agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga 4560
gaattatgca gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca 4620
acgatcggag gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact 4680
cgccttgatc gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc 4740
acgatgcctg tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact 4800
ctagcttccc ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt 4860
ctgcgctcgg cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt 4920
ggttctcgcg gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt 4980
atctacacga cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata 5040
ggtgcctcac tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag 5100
attgatttaa aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat 5160
ctcatgacca aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa 5220
aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca 5280
aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt 5340
ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgttcttct agtgtagccg 5400
tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatc 5460
ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga 5520
cgatagttac cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc 5580
agcttggagc gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc 5640
gccacgcttc ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca 5700

ggagagcgca cgagggagct tccagggga aacgcctggt atctttatag tcctgtcggg 5760
tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta 5820
tggaaaaacg ccagcaacgc ggccttttta cggttcctgg ccttttgctg gccttttgct 5880
cacat 5885

<210> 303

<211> 5885
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15794

<400> 303
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggtg gatctacgca cgctcggcta cagccagcag 480
caacaggaga agatcaaacc gaaggttcgt tcgacagtgg cgcagcacca cgaggcactg 540
gtcggccacg ggtttacaca cgcgcacatc gttgcgttaa gccaacaccc ggcagcgtta 600
gggaccgtcg ctgtcaagta tcaggacatg atcgcagcgt tgccagaggc gacacacgaa 660
gcgatcgttg gcgtcggcaa acagtggtcc ggcgcacgcg ctctggaggc cttgctcacg 720
gtggcgggag agttgagagg tccaccgtta cagttggaca caggccaact tctcaagatt 780
gcaaaacgtg gcggcgtgac cgcagtggag gcagtgcatg catggcgcaa tgcactgacg 840
ggtgccccgc tcaacttgac cccccagcag gtggtggcca tcgccagcaa taatggtggc 900
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 960
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 1020
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg 1080
gccatcgcca gcaatggcgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg 1140
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaataat 1200
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 1260
ggcttgaccc cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg 1320
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 1380
gtggtggcca tcgccagcaa taatggtggc aagcaggcgc tggagacggt ccagcggctg 1440
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 1500
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 1560
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 1620
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc 1680
cagcaggtgg tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag 1740
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 1800
gccagcaata atggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 1860
tgccaggccc acggcttgac cccccagcag gtggtggcca tcgccagcaa tggcggtggc 1920
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 1980
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 2040
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 2100
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 2160
gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg tggccatcgc cagcaatggc 2220
ggcggcaggc cggcgctgga gagcattgtt gcccagttat ctcgccctga tccgggatcc 2280
gacatcacca aatctaaaat ctctgaaaag atgaaaggcc agggcccatc cggaaatacc 2340
aaatataacg aagagttcct gctgtacctg gccggctttg tggacggtga cggtagcatc 2400
gttgctcaga ttaaaccaca tcagtcttgt aagtttaaac atcagctacg tttgaccttt 2460
aacgtgactc aaaagaccca gcgccgttgg tttctggaca aactagtgga tgaaattggc 2520
gttggtcatg tatacgattc tggtagcgtt tcctactacc agttaagcga aatcaagccg 2580
ctgcacaact tcctgactca actgcagccg tttctggaac tgaaacagaa acaggcaaac 2640
ctggttctga aaattatcga acagctgccg tctgcaaaag aatccccgga caaattcctg 2700
gaagtttgta cctgggtgga tcaggttgca gctctgaacg attctaagac gcgtaaaacc 2760
acttctgaaa cctgtcgtgc tgtgctggac agcctgagcg agaagaagaa atcctccccg 2820
gcggccggtg attcctctgt ttctaattcc gagcacattg ctcctctgtc tctgccttcc 2880
tctcctccat ctgttggttc taacaaaaaa ttcctgctgt atcttgctgg atttgtggat 2940
tctgatggct ccatcattgc tcagataaaa ccaaatcaat cttacaagtt caaacactac 3000
ctctccttga ccttttgtgt cactcagaag acacaaagaa ggtggttctt ggacaaattg 3060
gttgatcgta ttggtgtggg ctatgtccag gacagaggct ctgtgtcaga ctaccgtctg 3120

tctgaaatta agcctcttcg taactttctc acccaactgc aacccttctt gaagctcaaa 3180
cagaagcaag caaatctggt tttgaaaatc atcgagcaac tgccatctgc caaggagtcc 3240
cctgacaagt ttcttgaagt gtgtacttgg gtggatcagg ttgctgcctt gaatgactcc 3300
aagaccagaa aaaccacctc tgagactgtg agggcagttc tggatagcct ctctgagaag 3360
aaaaagtcct ctccttagta actcgagcct gcagcagctg aagctttgga cttctggcgc 3420
gcccaattcg ccctatagtg agtcgtatta cgtcgcgctc actggccgtc gttttacaac 3480
gtcgtgactg ggaaaaccct ggcgttaccc aacttaatcg ccttgcagca catccccctt 3540
tcgccagctg gcgtaatagc gaagaggccc gcaccgaaac gcccttccca acagttgcgc 3600
agcctgaatg gcgaatggga gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg 3660
ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct 3720
tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat cgggggctcc 3780
ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg 3840
atggttggcc tgtagtgggc catagccctg atagacggtt tttcgccctt tgacgttgga 3900
gtccacgttc tttaatagtg gactcttgtt ccaaactgga acaacactca accctatctc 3960
ggtctattct tttgatttat aagggatttt gccgatttcg gcctattggt taaaaaatga 4020
gctgatttaa caaaaattta acgcgaattt taacaaaata ttaacgctta caatttaggt 4080
ggcacttttc ggggaaatgt gcgcggaacc cctatttgtt tatttttcta aatacattca 4140
aatatgtatc cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg 4200
aagagtatga gtattcaaca tttccgtgtc gcccttattc cctttttgc ggcattttgc 4260
cttcctgttt ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg 4320
ggtgcacgag tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt 4380
cgccccgaag aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta 4440
ttatcccgta ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat 4500
gacttggttg agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga 4560
gaattatgca gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca 4620
acgatcggag gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact 4680
cgccttgatc gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc 4740
acgatgcctg tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact 4800
ctagcttccc ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt 4860
ctgcgctcgg cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt 4920
ggttctcgcg gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt 4980
atctacacga cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata 5040
ggtgcctcac tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag 5100
attgatttaa aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat 5160
ctcatgacca aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa 5220
aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca 5280
aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt 5340
ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgttcttct agtgtagccg 5400
tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatc 5460
ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga 5520
cgatagttac cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc 5580
agcttggagc gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc 5640
gccacgcttc ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca 5700
ggagagcgca cgagggagct tccaggggga aacgcctggt atctttatag tcctgtcggg 5760
tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta 5820
tggaaaaacg ccagcaacgc ggccttttta cggttcctgg ccttttgctg gccttttgct 5880
cacat 5885

<210> 304

<211> 10686
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS13449

<400> 304
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540

```
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt    600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg    660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc    720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa    780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag    840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg    900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga    960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag   1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa   1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc   1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca   1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg   1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat   1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc   1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg   1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg   1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt   1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca   1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata   1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac   1740
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa   1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag   1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg   1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta   1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc   2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg   2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa   2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt   2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca   2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc   2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata   2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg tggatctacg cacgctcggc   2460
tacagccagc agcaacagga gaagatcaaa ccgaaggttc gttcgacagt ggcgcagcac   2520
cacgaggcac tggtcggcca cgggtttaca cacgcgcaca tcgttgcgtt aagccaacac   2580
ccggcagcgt tagggaccgt cgctgtcaag tatcaggaca tgatcgcagc gttgccagag   2640
gcgacacacg aagcgatcgt tggcgtcggc aaacagtggt ccggcgcacg cgctctggag   2700
gccttgctca cggtggcggg agagttgaga ggtccaccgt tacagttgga cacaggccaa   2760
cttctcaaga ttgcaaaacg tggcggcgtg accgcagtgg aggcagtgca tgcatggcgc   2820
aatgcactga cgggtgcccc gctcaacttg acccccagc aggtggtggc catcgccagc   2880
aataatggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag   2940
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag   3000
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc   3060
cagcaggtgg tggccatcgc cagcaatggc ggtggcaagc aggcgctgga gacggtccag   3120
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc   3180
gccagcaata atggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg   3240
tgccaggccc acggcttgac cccccagcag gtggtggcca tcgccagcaa taatggtggc   3300
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg   3360
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg   3420
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg   3480
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg   3540
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagcaatatt   3600
ggtggcaagc aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac   3660
ggcttgaccc cccagcaggt ggtggccatc gccagcaata atggtggcaa gcaggcgctg   3720
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag   3780
gtggtggcca tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg   3840
ttgccggtgc tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc   3900
aataatggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag   3960
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag   4020
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg   4080
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag   4140
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc ctcagcaggt ggtggccatc   4200
gccagcaatg gcggcggcag gccggcgctg gagagcattg ttgcccagtt atctcgccct   4260
gatccgggat ccgacatcac caaatctaaa atctctgaaa agatgaaagg ccagggccca   4320
tccggaaata ccaaatataa cgaagagttc ctgctgtacc tggccggctt tgtggacggt   4380
gacggtagca tcgttgctca gattaaacca catcagtctt gtaagtttaa acatcagcta   4440
cgtttgacct ttaacgtgac tcaaaagacc cagcgccgtt ggtttctgga caaactagtg   4500
gatgaaattg gcgttggtca tgtatacgat tctggtagcg tttcctacta ccagttaagc   4560
gaaatcaagc cgctgcacaa cttcctgact caactgcagc cgtttctgga actgaaacag   4620
```

aaacaggcaa acctggttct gaaaattatc gaacagctgc cgtctgcaaa agaatccccg 4680
gacaaattcc tggaagtttg tacctgggtg gatcaggttg cagctctgaa cgattctaag 4740
acgcgtaaaa ccacttctga aaccgttcgt gctgtgctgg acagcctgag cgagaagaag 4800
aaatcctccc cggcggccga ctgataactc gagcgatcct ctagacgagc tcctcgagcc 4860
tgcagcagct gaagctttgg acttcttcgc cagaggtttg gtcaagtctc caatcaaggt 4920
tgtcggcttg tctaccttgc cagaaattta cgaaaagatg gaaaagggtc aaatcgttgg 4980
tagatacgtt gttgacactt ctaaataagc gaatttctta tgatttatga tttttattat 5040
taaataagtt ataaaaaaaa taagtgtata caaattttaa agtgactctt aggttttaaa 5100
acgaaaattc ttattcttga gtaactcttt cctgtaggtc aggttgcttt ctcaggtata 5160
gcatgaggtc gctcttattg accacacctc taccggcatg caagcttggc gtaatcatgg 5220
tcatagctgt ttcctgtgtg aaattgttat ccgctcacaa ttccacacaa catacgagcc 5280
ggaagcataa agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaattgcg 5340
ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt gccagcagat ctattacatt 5400
atgggtggta tgttggaata aaaatcaact atcatctact aactagtatt tacgttacta 5460
gtatattatc atatacggtg ttagaagatg acgcaaatga tgagaaatag tcatctaaat 5520
tagtggaagc tgaaacgcaa ggattgataa tgtaatagga tcaatgaata ttaacatata 5580
aaatgatgat aataatattt atagaattgt gtagaattgc agattccctt ttatggattc 5640
ctaaatcctc gaggagaact tctagtatat ctacatacct aatattattg ccttattaaa 5700
aatggaatcc caacaattac atcaaaatcc acattctctt caaaatcaat tgtcctgtac 5760
ttccttgttc atgtgtgttc aaaaacgtta tatttatagg ataattatac tctatttctc 5820
aacaagtaat tggttgtttg gccgagcggt ctaaggcgcc tgattcaaga aatatcttga 5880
ccgcagttaa ctgtgggaat actcaggtat cgtaagatgc aagagttcga atctcttagc 5940
aaccattatt tttttcctca acataacgag aacacacagg ggcgctatcg cacagaatca 6000
aattcgatga ctggaaattt tttgttaatt tcagaggtcg cctgacgcat atacctttt 6060
caactgaaaa attgggagaa aaaggaaagg tgagagccgc ggaaccggct tttcatatag 6120
aatagagaag cgttcatgac taaatgcttg catcacaata cttgaagttg acaatattat 6180
ttaaggacct attgtttttt ccaataggtg gttagcaatc gtcttacttt ctaacttttc 6240
ttacctttta catttcagca atatatatat atatatttca aggatatacc attctaatgt 6300
ctgcccctaa gaagatcgtc gttttgccag gtgaccacgt tggtcaagaa atcacagccg 6360
aagccattaa ggttcttaaa gctatttctg atgttcgttc caatgtcaag ttcgatttcg 6420
aaaatcattt aattggtggt gctgctatcg atgctacagg tgtcccactt ccagatgagg 6480
cgctggaagc ctccaagaag gttgatgccg ttttgttagg tgctgtgggt ggtcctaaat 6540
ggggtaccgg tagtgttaga cctgaacaag gtttactaaa aatccgtaaa gaacttcaat 6600
tgtacgccaa cttaagacca tgtaactttg catccgactc tcttttagac ttatctccaa 6660
tcaagccaca atttgctaaa ggtactgact tcgttgttgt cagagaatta gtgggaggta 6720
tttactttgg taagagaaag gaagacgatg gtgatggtgt cgcttgggat agtgaacaat 6780
acaccgttcc agaagtgcaa agaatcacaa gaatggccgc tttcatggcc ctacaacatg 6840
agccaccatt gcctatttgg tccttggata aagctaatgt tttggcctct tcaagattat 6900
ggagaaaaac tgtggaggaa accatcaaga acgaattccc tacattgaag gttcaacatc 6960
aattgattga ttctgccgcc atgatcctag ttaagaaccc aacccaccta aatggtatta 7020
taatcaccag caacatgttt ggtgatatca tctccgatga agcctccgtt atcccaggtt 7080
ccttgggttt gttgccatct gcgtccttgg cctctttgcc agacaagaac accgcatttg 7140
gtttgtacga accatgccac ggttctgctc cagatttgcc aaagaataag gtcaacccta 7200
tcgccactat cttgtctgct gcaatgatgt tgaaattgtc attgaacttg cctgaagaag 7260
gtaaggccat tgaagatgca gttaaaaagg ttttggatgc aggtatcaga actggtgatt 7320
taggtggttc caacagtacc acggaagtcg gtgatgctgt cgccgaagaa gttaagaaaa 7380
tccttgctta aaaagattct ctttttttat gatatttgta cataaacttt ataaatgaaa 7440
ttcataatag aaacgacacg aaattacaaa atggaatatg ttcatagggt agacgaaact 7500
atatacgcaa tctacataca tttatcaaga aggagaaaaa ggaggatgta aaggaataca 7560
ggtaagcaaa ttgatactaa tggctcaacg tgataaggaa aaagaattgc actttaacat 7620
taatattgac aaggaggagg gcaccacaca aaaagttagg tgtaacagaa aatcatgaaa 7680
ctatgattcc taatttatat attggaggat tttctctaaa aaaaaaaaaa tacaacaaat 7740
aaaaaacact caatgacctg accatttgat ggagtttaag tcaatacctt cttgaaccat 7800
ttcccataat ggtgaaagtt ccctcaagaa ttttactctg tcagaaacgg ccttaacgac 7860
gtagtcgacc tcctcttcag tactaaatct accaatacca aatctgatgg aagaatgggc 7920
taatgcatca tccttaccca gcgcatgtaa aacataagaa ggttctaggg aagcagatgt 7980
acaggctgaa cccgaggata atgcgatatc ccttagtgcc atcaataaag attctccttc 8040
cacgtaggcg aaagaaacgt taacacaccc tggataacga tgatctggag atccgttcaa 8100
cgtggtatgt tcagcggata atagaccttt gactaattta tcggatagtc ttttgatgtg 8160
agcttggtcg ttgtcaaatt ctttcttcat caatctcgca gcttcaccaa atcccgctac 8220
caatgggggg gccaaagtac cagatctgct gcattaatga atcggccaac gcgcggggag 8280
aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt 8340
cgttcggctg cggcgagcgg tatcagcatc gatgaattcc acggactata gactatacta 8400
gtatactccg tctactgtac gatacacttc cgctcaggtc cttgtccttt aacgaggcct 8460
taccactctt ttgttactct attgatccag ctcagcaaag gcagtgtgat ctaagattct 8520
atcttcgcga tgtagtaaaa ctagctagac cgagaaagag actagaaatg caaaaggcac 8580
ttctacaatg gctgccatca ttattatccg atgtgacgct gcagcttctc aatgatattc 8640
gaatacgctt tgaggagata cagcctaata tccgacaaac tgttttacag atttacgatc 8700

gtacttgtta cccatcattg aattttgaac atccgaacct gggagttttc cctgaaacag 8760
atagtatatt tgaacctgta taataatata tagtctagcg ctttacggaa gacaatgtat 8820
gtatttcggt tcctggagaa actattgcat ctattgcata ggtaatcttg cacgtcgcat 8880
ccccggttca ttttctgcgt ttccatcttg cacttcaata gcatatcttt gttaacgaag 8940
catctgtgct tcattttgta gaacaaaaat gcaacgcgag agcgctaatt tttcaaacaa 9000
agaatctgag ctgcattttt acagaacaga aatgcaacgc gaaagcgcta ttttaccaac 9060
gaagaatctg tgcttcattt ttgtaaaaca aaaatgcaac gcgagagcgc taatttttca 9120
aacaaagaat ctgagctgca tttttacaga acagaaatgc aacgcgagag cgctatttta 9180
ccaacaaaga atctatactt ctttttttgtt ctacaaaaat gcatcccgag agcgctattt 9240
ttctaacaaa gcatcttaga ttacttttttt tctcctttgt gcgctctata atgcagtctc 9300
ttgataactt tttgcactgt aggtccgtta aggttagaag aaggctactt tggtgtctat 9360
tttctcttcc ataaaaaaag cctgactcca cttcccgcgt ttactgatta ctagcgaagc 9420
tgcgggtgca ttttttcaag ataaaggcat ccccgattat attctatacc gatgtggatt 9480
gcgcatactt tgtgaacaga aagtgatagc gttgatgatt cttcattggt cagaaaatta 9540
tgaacggttt cttctatttt gtctctatat actacgtata ggaaatgttt acattttcgt 9600
attgttttcg attcactcta tgaatagttc ttactacaat ttttttgtct aaagagtaat 9660
actagagata aacataaaaa atgtagaggt cgagtttaga tgcaagttca aggagcgaaa 9720
ggtggatggg taggttatat agggatatag cacagagata tatagcaaag agatactttt 9780
gagcaatgtt tgtggaagcg gtattcgcaa tattttagta gctcgttaca gtccggtgcg 9840
tttttggttt tttgaaagtg cgtcttcaga gcgcttttgg ttttcaaaag cgctctgaag 9900
ttcctatact ttctagagaa taggaacttc ggaataggaa cttcaaagcg tttccgaaaa 9960
cgagcgcttc cgaaaatgca acgcgagctg cgcacataca gctcactgtt cacgtcgcac 10020
ctatatctgc gtgttgcctg tatatatata tacatgagaa gaacggcata gtgcgtgttt 10080
atgcttaaat gcgtacttat atgcgtctat ttatgtagga tgaaaggtag tctagtacct 10140
cctgtgatat tatcccattc catgcggggt atcgtatgct tccttcagca ctacccttta 10200
gctgttctat atgctgccac tcctcaattg gattagtctc atccttcaat gctatcattt 10260
cctttgatat tggatcatat gcatagtacc gagaaactag tgcgaagtag tgatcaggta 10320
ttgctgttat ctgatgagta tacgttgtcc tggccacggc agaagcacgc ttatcgctcc 10380
aatttcccac aacattagtc aactccgtta ggcccttcat tgaaagaaat gaggtcatca 10440
aatgtcttcc aatgtgagat tttgggccat tttttatagc aaagattgaa taaggcgcat 10500
ttttcttcaa agctttattg tacgatctga ctaagttatc ttttaataat tggtattcct 10560
gtttattgct tgaagaattg ccggtcctat ttactcgttt taggactggt tcagaattca 10620
tcgatgctca ctcaaaggtc ggtaatacgg ttatccacag aatcagggga taacgcagga 10680
aagaac 10686

<210> 305

<211> 10686
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS13450

<400> 305
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320

```
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc   1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg   1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg   1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt   1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca   1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata   1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac   1740
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa   1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag   1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg   1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta   1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc   2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg   2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa   2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt   2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca   2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc   2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata   2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg tggatctacg cacgctcggc   2460
tacagccagc agcaacagga gaagatcaaa ccgaaggttc gttcgacagt ggcgcagcac   2520
cacgaggcac tggtcggcca cgggtttaca cacgcgcaca tcgttgcgtt aagccaacac   2580
ccggcagcgt tagggaccgt cgctgtcaag tatcaggaca tgatcgcagc gttgccagag   2640
gcgacacacg aagcgatcgt tggcgtcggc aaacagtggt ccggcgcacg cgctctggag   2700
gccttgctca cggtggcggg agagttgaga ggtccaccgt tacagttgga cacaggccaa   2760
cttctcaaga ttgcaaaacg tggcggcgtg accgcagtgg aggcagtgca tgcatggcgc   2820
aatgcactga cgggtgcccc gctcaacttg acccccccagc aggtggtggc catcgccagc   2880
aataatggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag   2940
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag   3000
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc   3060
cagcaggtgg tggccatcgc cagcaatggc ggtggcaagc aggcgctgga gacggtccag   3120
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc   3180
gccagcaata atggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg   3240
tgccaggccc acggcttgac cccccagcag gtggtggcca tcgccagcaa tggcggtggc   3300
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg   3360
acccccccagc aggtggtggc catcgccagc aataatggtg gcaagcaggc gctggagacg   3420
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg   3480
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg   3540
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc   3600
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac   3660
ggcttgaccc cccagcaggt ggtggccatc gccagcaata atggtggcaa gcaggcgctg   3720
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag   3780
gtggtggcca tcgccagcaa taatggtggc aagcaggcgc tggagacggt ccagcggctg   3840
ttgccggtgc tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc   3900
aatggcggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag   3960
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag   4020
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg   4080
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag   4140
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc ctcagcaggt ggtggccatc   4200
gccagcaatg gcggcggcag gccggcgctg gagagcattg ttgcccagtt atctcgccct   4260
gatccgggat ccgacatcac caaatctaaa atctctgaaa agatgaaagg ccagggccca   4320
tccggaaata ccaaatataa cgaagagttc ctgctgtacc tggccggctt tgtggacggt   4380
gacggtagca tcgttgctca gattaaacca catcagtctt gtaagtttaa acatcagcta   4440
cgtttgacct ttaacgtgac tcaaaagacc cagcgccgtt ggtttctgga caaactagtg   4500
gatgaaattg gcgttggtca tgtatacgat tctggtagcg tttcctacta ccagttaagc   4560
gaaatcaagc cgctgcacaa cttcctgact caactgcagc cgtttctgga actgaaacag   4620
aaacaggcaa acctggttct gaaaattatc gaacagctgc cgtctgcaaa agaatccccg   4680
gacaaattcc tggaagtttg tacctgggtg gatcaggttg cagctctgaa cgattctaag   4740
acgcgtaaaa ccacttctga aaccgttcgt gctgtgctgg acagcctgag cgagaagaag   4800
aaatcctccc cggcggccga ctgataactc gagcgatcct ctagacgagc tcctcgagcc   4860
tgcagcagct gaagctttgg acttcttcgc cagaggtttg gtcaagtctc caatcaaggt   4920
tgtcggcttg tctaccttgc cagaaattta cgaaaagatg gaaaagggtc aaatcgttgg   4980
tagatacgtt gttgacactt ctaaataagc gaatttctta tgatttatga tttttattat   5040
taaataagtt ataaaaaaaa taagtgtata caaattttaa agtgactctt aggttttaaa   5100
acgaaaattc ttattcttga gtaactcttt cctgtaggtc aggttgcttt ctcaggtata   5160
gcatgaggtc gctcttattg accacacctc taccggcatg caagcttggc gtaatcatgg   5220
tcatagctgt ttcctgtgtg aaattgttat ccgctcacaa ttccacacaa catacgagcc   5280
ggaagcataa agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaattgcg   5340
ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt gccagcagat ctattacatt   5400
```

atgggtggta tgttggaata aaaatcaact atcatctact aactagtatt tacgttacta 5460
gtatattatc atatacggtg ttagaagatg acgcaaatga tgagaaatag tcatctaaat 5520
tagtggaagc tgaaacgcaa ggattgataa tgtaatagga tcaatgaata ttaacatata 5580
aaatgatgat aataatattt atagaattgt gtagaattgc agattccctt ttatggattc 5640
ctaaatcctc gaggagaact tctagtatat ctacatacct aatattattg ccttattaaa 5700
aatggaatcc caacaattac atcaaaatcc acattctctt caaaatcaat tgtcctgtac 5760
ttccttgttc atgtgtgttc aaaaacgtta tatttatagg ataattatac tctatttctc 5820
aacaagtaat tggttgtttg gccgagcggt ctaaggcgcc tgattcaaga aatatcttga 5880
ccgcagttaa ctgtgggaat actcaggtat cgtaagatgc aagagttcga atctcttagc 5940
aaccattatt tttttcctca acataacgag aacacacagg ggcgctatcg cacagaatca 6000
aattcgatga ctggaaattt tttgttaatt tcagaggtcg cctgacgcat atacctttt 6060
caactgaaaa attgggagaa aaaggaaagg tgagagccgc ggaaccggct tttcatatag 6120
aatagagaag cgttcatgac taaatgcttg catcacaata cttgaagttg acaatattat 6180
ttaaggacct attgtttttt ccaataggtg gttagcaatc gtcttacttt ctaacttttc 6240
ttacctttta catttcagca atatatatat atatatttca aggatatacc attctaatgt 6300
ctgcccctaa gaagatcgtc gttttgccag gtgaccacgt tggtcaagaa atcacagccg 6360
aagccattaa ggttcttaaa gctatttctg atgttcgttc caatgtcaag ttcgatttcg 6420
aaaatcattt aattggtggt gctgctatcg atgctacagg tgtcccactt ccagatgagg 6480
cgctggaagc ctccaagaag gttgatgccg ttttgttagg tgctgtgggt ggtcctaaat 6540
ggggtaccgg tagtgttaga cctgaacaag gtttactaaa aatccgtaaa gaacttcaat 6600
tgtacgccaa cttaagacca tgtaactttg catccgactc tcttttagac ttatctccaa 6660
tcaagccaca atttgctaaa ggtactgact tcgttgttgt cagagaatta gtgggaggta 6720
tttactttgg taagagaaag gaagacgatg gtgatggtgt cgcttgggat agtgaacaat 6780
acaccgttcc agaagtgcaa agaatcacaa gaatggccgc tttcatggcc ctacaacatg 6840
agccaccatt gcctatttgg tccttggata aagctaatgt tttggcctct tcaagattat 6900
ggagaaaaac tgtggaggaa accatcaaga acgaattccc tacattgaag gttcaacatc 6960
aattgattga ttctgccgcc atgatcctag ttaagaaccc aacccaccta aatggtatta 7020
taatcaccag caacatgttt ggtgatatca tctccgatga agcctccgtt atcccaggtt 7080
ccttgggttt gttgccatct gcgtccttgg cctctttgcc agacaagaac accgcatttg 7140
gtttgtacga accatgccac ggttctgctc cagatttgcc aaagaataag gtcaacccta 7200
tcgccactat cttgtctgct gcaatgatgt tgaaattgtc attgaacttg cctgaagaag 7260
gtaaggccat tgaagatgca gttaaaaagg ttttggatgc aggtatcaga actggtgatt 7320
taggtggttc caacagtacc acggaagtcg gtgatgctgt cgccgaagaa gttaagaaaa 7380
tccttgctta aaaagattct ctttttttat gatatttgta cataaacttt ataaatgaaa 7440
ttcataatag aaacgacacg aaattacaaa atggaatatg ttcatagggt agacgaaact 7500
atatacgcaa tctacataca tttatcaaga aggagaaaaa ggaggatgta aaggaataca 7560
ggtaagcaaa ttgatactaa tggctcaacg tgataaggaa aaagaattgc actttaacat 7620
taatattgac aaggaggagg gcaccacaca aaaagttagg tgtaacagaa aatcatgaaa 7680
ctatgattcc taatttatat attggaggat tttctctaaa aaaaaaaaaa tacaacaaat 7740
aaaaaacact caatgacctg accatttgat ggagtttaag tcaatacctt cttgaaccat 7800
ttcccataat ggtgaaagtt ccctcaagaa ttttactctg tcagaaacgg ccttaacgac 7860
gtagtcgacc tcctcttcag tactaaatct accaatacca aatctgatgg aagaatgggc 7920
taatgcatca tccttaccca gcgcatgtaa aacataagaa ggttctaggg aagcagatgt 7980
acaggctgaa cccgaggata atgcgatatc ccttagtgcc atcaataaag attctccttc 8040
cacgtaggcg aaagaaacgt taacacaccc tggataacga tgatctggag atccgttcaa 8100
cgtggtatgt tcagcggata atagaccttt gactaattta tcggatagtc ttttgatgtg 8160
agcttggtcg ttgtcaaatt ctttcttcat caatctcgca gcttcaccaa atcccgctac 8220
caatgggggg gccaaagtac cagatctgct gcattaatga atcggccaac gcgcggggag 8280
aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt 8340
cgttcggctg cggcgagcgg tatcagcatc gatgaattcc acggactata gactatacta 8400
gtatactccg tctactgtac gatacacttc cgctcaggtc cttgtccttt aacgaggcct 8460
taccactctt ttgttactct attgatccag ctcagcaaag gcagtgtgat ctaagattct 8520
atcttcgcga tgtagtaaaa ctagctagac cgagaaagag actagaaatg caaaaggcac 8580
ttctacaatg gctgccatca ttattatccg atgtgacgct gcagcttctc aatgatattc 8640
gaatacgctt tgaggagata cagcctaata tccgacaaac tgttttacag atttacgatc 8700
gtacttgtta cccatcattg aattttgaac atccgaacct gggagtttc cctgaaacag 8760
atagtatatt tgaacctgta taataatata tagtctagcg ctttacggaa gacaatgtat 8820
gtatttcggt tcctggagaa actattgcat ctattgcata ggtaatcttg cacgtcgcat 8880
ccccggttca ttttctgcgt ttccatcttg cacttcaata gcatatcttt gttaacgaag 8940
catctgtgct tcattttgta gaacaaaaat gcaacgcgag agcgctaatt tttcaaacaa 9000
agaatctgag ctgcattttt acagaacaga aatgcaacgc gaaagcgcta ttttaccaac 9060
gaagaatctg tgcttcattt ttgtaaaaca aaaatgcaac gcgagagcgc taatttttca 9120
aacaaagaat ctgagctgca tttttacaga acagaaatgc aacgcgagag cgctatttta 9180
ccaacaaaga atctatactt ctttttttgtt ctacaaaaat gcatcccgag agcgctattt 9240
ttctaacaaa gcatcttaga ttacttttt tctcctttgt gcgctctata atgcagtctc 9300
ttgataactt tttgcactgt aggtccgtta aggttagaag aaggctactt tggtgtctat 9360
tttctcttcc ataaaaaaag cctgactcca cttcccgcgt ttactgatta ctagcgaagc 9420
tgcgggtgca ttttttcaag ataaaggcat ccccgattat attctatacc gatgtggatt 9480

gcgcatactt tgtgaacaga aagtgatagc gttgatgatt cttcattggt cagaaaatta 9540
tgaacggttt cttctatttt gtctctatat actacgtata ggaaatgttt acattttcgt 9600
attgttttcg attcactcta tgaatagttc ttactacaat ttttttgtct aaagagtaat 9660
actagagata aacataaaaa atgtagaggt cgagtttaga tgcaagttca aggagcgaaa 9720
ggtggatggg taggttatat agggatatag cacagagata tatagcaaag agatactttt 9780
gagcaatgtt tgtggaagcg gtattcgcaa tattttagta gctcgttaca gtccggtgcg 9840
ttttggtttt tttgaaagtg cgtcttcaga gcgcttttgg ttttcaaaag cgctctgaag 9900
ttcctatact ttctagagaa taggaacttc ggaataggaa cttcaaagcg tttccgaaaa 9960
cgagcgcttc cgaaaatgca acgcgagctg cgcacataca gctcactgtt cacgtcgcac 10020
ctatatctgc gtgttgcctg tatatatata tacatgagaa gaacggcata gtgcgtgttt 10080
atgcttaaat gcgtacttat atgcgtctat ttatgtagga tgaaaggtag tctagtacct 10140
cctgtgatat tatcccattc catgcggggt atcgtatgct tccttcagca ctacccttta 10200
gctgttctat atgctgccac tcctcaattg gattagtctc atccttcaat gctatcattt 10260
cctttgatat tggatcatat gcatagtacc gagaaactag tgcgaagtag tgatcaggta 10320
ttgctgttat ctgatgagta tacgttgtcc tggccacggc agaagcacgc ttatcgctcc 10380
aatttcccac aacattagtc aactccgtta ggcccttcat tgaaagaaat gaggtcatca 10440
aatgtcttcc aatgtgagat tttgggccat tttttatagc aaagattgaa taaggcgcat 10500
ttttcttcaa agctttattg tacgatctga ctaagttatc ttttaataat tggtattcct 10560
gtttattgct tgaagaattg ccggtcctat ttactcgttt taggactggt tcagaattca 10620
tcgatgctca ctcaaaggtc ggtaatacgg ttatccacag aatcagggga taacgcagga 10680
aagaac 10686

<210> 306

<211> 10686
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS13451

<400> 306
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2100

gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg tggatctacg cacgctcggc 2460
tacagccagc agcaacagga gaagatcaaa ccgaaggttc gttcgacagt ggcgcagcac 2520
cacgaggcac tggtcggcca cgggtttaca cacgcgcaca tcgttgcgtt aagccaacac 2580
ccggcagcgt tagggaccgt cgctgtcaag tatcaggaca tgatcgcagc gttgccagag 2640
gcgacacacg aagcgatcgt tggcgtcggc aaacagtggt ccggcgcacg cgctctggag 2700
gccttgctca cggtggcggg agagttgaga ggtccaccgt tacagttgga cacaggccaa 2760
cttctcaaga ttgcaaaacg tggcggcgtg accgcagtgg aggcagtgca tgcatggcgc 2820
aatgcactga cgggtgcccc gctcaacttg acccccccagc aggtggtggc catcgccagc 2880
aataatggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2940
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 3000
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 3060
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 3120
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 3180
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3240
tgccaggccc acggcttgac ccccccagcag gtggtggcca tcgccagcaa taatggtggc 3300
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 3360
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 3420
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg 3480
gccatcgcca gcaatggcgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg 3540
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaataat 3600
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 3660
ggcttgaccc cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg 3720
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 3780
gtggtggcca tcgccagcaa taatggtggc aagcaggcgc tggagacggt ccagcggctg 3840
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 3900
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 3960
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 4020
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc 4080
cagcaggtgg tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag 4140
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc ctcagcaggt ggtggccatc 4200
gccagcaatg gcggcggcag gccggcgctg gagagcattg ttgcccagtt atctcgccct 4260
gatccgggat ccgacatcac caaatctaaa atctctgaaa agatgaaagg ccagggccca 4320
tccggaaata ccaaatataa cgaagagttc ctgctgtacc tggccggctt tgtggacggt 4380
gacggtagca tcgttgctca gattaaacca catcagtctt gtaagtttaa acatcagcta 4440
cgtttgacct ttaacgtgac tcaaaagacc cagcgccgtt ggtttctgga caaactagtg 4500
gatgaaattg gcgttggtca tgtatacgat tctggtagcg tttcctacta ccagttaagc 4560
gaaatcaagc cgctgcacaa cttcctgact caactgcagc cgtttctgga actgaaacag 4620
aaacaggcaa acctggttct gaaaattatc gaacagctgc cgtctgcaaa agaatccccg 4680
gacaaattcc tggaagtttg tacctgggtg gatcaggttg cagctctgaa cgattctaag 4740
acgcgtaaaa ccacttctga aaccgttcgt gctgtgctgg acagcctgag cgagaagaag 4800
aaatcctccc cggcggccga ctgataactc gagcgatcct ctagacgagc tcctcgagcc 4860
tgcagcagct gaagctttgg acttcttcgc cagaggtttg gtcaagtctc caatcaaggt 4920
tgtcggcttg tctaccttgc cagaaattta cgaaaagatg gaaaagggtc aaatcgttgg 4980
tagatacgtt gttgacactt ctaaataagc gaatttctta tgatttatga tttttattat 5040
taaataagtt ataaaaaaaa taagtgtata caaattttaa agtgactctt aggttttaaa 5100
acgaaaattc ttattcttga gtaactcttt cctgtaggtc aggttgcttt ctcaggtata 5160
gcatgaggtc gctcttattg accacacctc taccggcatg caagcttggc gtaatcatgg 5220
tcatagctgt ttcctgtgtg aaattgttat ccgctcacaa ttccacacaa catacgagcc 5280
ggaagcataa agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaattgcg 5340
ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt gccagcagat ctattacatt 5400
atgggtggta tgttggaata aaaatcaact atcatctact aactagtatt tacgttacta 5460
gtatattatc atatacggtg ttagaagatg acgcaaatga tgagaaatag tcatctaaat 5520
tagtggaagc tgaaacgcaa ggattgataa tgtaatagga tcaatgaata ttaacatata 5580
aaatgatgat aataatattt atagaattgt gtagaattgc agattccctt ttatggattc 5640
ctaaatcctc gaggagaact tctagtatat ctacatacct aatattattg ccttattaaa 5700
aatggaatcc caacaattac atcaaaatcc acattctctt caaaatcaat tgtcctgtac 5760
ttccttgttc atgtgtgttc aaaaacgtta tatttatagg ataattatac tctatttctc 5820
aacaagtaat tggttgtttg gccgagcggt ctaaggcgcc tgattcaaga aatatcttga 5880
ccgcagttaa ctgtgggaat actcaggtat cgtaagatgc aagagttcga atctcttagc 5940
aaccattatt tttttcctca acataacgag aacacacagg ggcgctatcg cacagaatca 6000
aattcgatga ctggaaattt tttgttaatt tcagaggtcg cctgacgcat atacctttt 6060
caactgaaaa attgggagaa aaaggaaagg tgagagccgc ggaaccggct tttcatatag 6120
aatagagaag cgttcatgac taaatgcttg catcacaata cttgaagttg acaatattat 6180

ttaaggacct attgtttttt ccaataggtg gttagcaatc gtcttacttt ctaacttttc 6240
ttacctttta catttcagca atatatatat atatatttca aggatatacc attctaatgt 6300
ctgcccctaa gaagatcgtc gttttgccag gtgaccacgt tggtcaagaa atcacagccg 6360
aagccattaa ggttcttaaa gctatttctg atgttcgttc caatgtcaag ttcgatttcg 6420
aaaatcattt aattggtggt gctgctatcg atgctacagg tgtcccactt ccagatgagg 6480
cgctggaagc ctccaagaag gttgatgccg ttttgttagg tgctgtgggt ggtcctaaat 6540
ggggtaccgg tagtgttaga cctgaacaag gtttactaaa aatccgtaaa gaacttcaat 6600
tgtacgccaa cttaagacca tgtaactttg catccgactc tcttttagac ttatctccaa 6660
tcaagccaca atttgctaaa ggtactgact tcgttgttgt cagagaatta gtgggaggta 6720
tttactttgg taagagaaag gaagacgatg gtgatggtgt cgcttgggat agtgaacaat 6780
acaccgttcc agaagtgcaa agaatcacaa gaatggccgc tttcatggcc ctacaacatg 6840
agccaccatt gcctatttgg tccttggata aagctaatgt tttggcctct tcaagattat 6900
ggagaaaaac tgtggaggaa accatcaaga acgaattccc tacattgaag gttcaacatc 6960
aattgattga ttctgccgcc atgatcctag ttaagaaccc aacccaccta aatggtatta 7020
taatcaccag caacatgttt ggtgatatca tctccgatga agcctccgtt atcccaggtt 7080
ccttgggttt gttgccatct gcgtccttgg cctctttgcc agacaagaac accgcatttg 7140
gtttgtacga accatgccac ggttctgctc cagatttgcc aaagaataag gtcaacccta 7200
tcgccactat cttgtctgct gcaatgatgt tgaaattgtc attgaacttg cctgaagaag 7260
gtaaggccat tgaagatgca gttaaaaagg ttttggatgc aggtatcaga actggtgatt 7320
taggtggttc caacagtacc acggaagtcg gtgatgctgt cgccgaagaa gttaagaaaa 7380
tccttgctta aaaagattct ctttttttat gatatttgta cataaacttt ataaatgaaa 7440
ttcataatag aaacgacacg aaattacaaa atggaatatg ttcatagggt agacgaaact 7500
atatacgcaa tctacataca tttatcaaga aggagaaaaa ggaggatgta aaggaataca 7560
ggtaagcaaa ttgatactaa tggctcaacg tgataaggaa aaagaattgc actttaacat 7620
taatattgac aaggaggagg gcaccacaca aaaagttagg tgtaacagaa aatcatgaaa 7680
ctatgattcc taatttatat attggaggat tttctctaaa aaaaaaaaaa tacaacaaat 7740
aaaaaacact caatgacctg accatttgat ggagtttaag tcaatacctt cttgaaccat 7800
ttcccataat ggtgaaagtt ccctcaagaa ttttactctg tcagaaacgg ccttaacgac 7860
gtagtcgacc tcctcttcag tactaaatct accaatacca aatctgatgg aagaatgggc 7920
taatgcatca tccttaccca gcgcatgtaa aacataagaa ggttctaggg aagcagatgt 7980
acaggctgaa cccgaggata atgcgatatc ccttagtgcc atcaataaag attctccttc 8040
cacgtaggcg aaagaaacgt taacacaccc tggataacga tgatctggag atccgttcaa 8100
cgtggtatgt tcagcggata atagaccttt gactaattta tcggatagtc ttttgatgtg 8160
agcttggtcg ttgtcaaatt ctttcttcat caatctcgca gcttcaccaa atcccgctac 8220
caatgggggg gccaaagtac cagatctgct gcattaatga atcggccaac gcgcggggag 8280
aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt 8340
cgttcggctg cggcgagcgg tatcagcatc gatgaattcc acggactata gactatacta 8400
gtatactccg tctactgtac gatacacttc cgctcaggtc cttgtccttt aacgaggcct 8460
taccactctt ttgttactct attgatccag ctcagcaaag gcagtgtgat ctaagattct 8520
atcttcgcga tgtagtaaaa ctagctagac cgagaaagag actagaaatg caaaaggcac 8580
ttctacaatg gctgccatca ttattatccg atgtgacgct gcagcttctc aatgatattc 8640
gaatacgctt tgaggagata cagcctaata tccgacaaac tgttttacag atttacgatc 8700
gtacttgtta cccatcattg aattttgaac atccgaacct gggagttttc cctgaaacag 8760
atagtatatt tgaacctgta taataatata tagtctagcg ctttacggaa gacaatgtat 8820
gtatttcggt tcctggagaa actattgcat ctattgcata ggtaatcttg cacgtcgcat 8880
ccccggttca ttttctgcgt ttccatcttg cacttcaata gcatatcttt gttaacgaag 8940
catctgtgct tcattttgta gaacaaaaat gcaacgcgag agcgctaatt tttcaaacaa 9000
agaatctgag ctgcattttt acagaacaga aatgcaacgc gaaagcgcta ttttaccaac 9060
gaagaatctg tgcttcattt ttgtaaaaca aaaatgcaac gcgagagcgc taatttttca 9120
aacaaagaat ctgagctgca tttttacaga acagaaatgc aacgcgagag cgctatttta 9180
ccaacaaaga atctatactt ctttttttgtt ctacaaaaat gcatcccgag agcgctattt 9240
ttctaacaaa gcatcttaga ttacttttt tctcctttgt gcgctctata atgcagtctc 9300
ttgataactt tttgcactgt aggtccgtta aggttagaag aaggctactt tggtgtctat 9360
tttctcttcc ataaaaaaag cctgactcca cttcccgcgt ttactgatta ctagcgaagc 9420
tgcgggtgca ttttttcaag ataaaggcat ccccgattat attctatacc gatgtggatt 9480
gcgcatactt tgtgaacaga aagtgatagc gttgatgatt cttcattggt cagaaaatta 9540
tgaacggttt cttctatttt gtctctatat actacgtata ggaaatgttt acattttcgt 9600
attgttttcg attcactcta tgaatagttc ttactacaat ttttttgtct aaagagtaat 9660
actagagata aacataaaaa atgtagaggt cgagtttaga tgcaagttca aggagcgaaa 9720
ggtggatggg taggttatat agggatatag cacagagata tatagcaaag agatactttt 9780
gagcaatgtt tgtggaagcg gtattcgcaa tattttagta gctcgttaca gtccggtgcg 9840
tttttggttt tttgaaagtg cgtcttcaga gcgcttttgg ttttcaaaag cgctctgaag 9900
ttcctatact ttctagagaa taggaacttc ggaataggaa cttcaaagcg tttccgaaaa 9960
cgagcgcttc cgaaaatgca acgcgagctg cgcacataca gctcactgtt cacgtcgcac 10020
ctatatctgc gtgttgcctg tatatatata tacatgagaa gaacggcata gtgcgtgttt 10080
atgcttaaat gcgtacttat atgcgtctat ttatgtagga tgaaaggtag tctagtacct 10140
cctgtgatat tatcccattc catgcggggt atcgtatgct tccttcagca ctacccttta 10200
gctgttctat atgctgccac tcctcaattg gattagtctc atccttcaat gctatcattt 10260

cctttgatat tggatcatat gcatagtacc gagaaactag tgcgaagtag tgatcaggta 10320
ttgctgttat ctgatgagta tacgttgtcc tggccacggc agaagcacgc ttatcgctcc 10380
aatttcccac aacattagtc aactccgtta ggcccttcat tgaaagaaat gaggtcatca 10440
aatgtcttcc aatgtgagat tttgggccat tttttatagc aaagattgaa taaggcgcat 10500
ttttcttcaa agctttattg tacgatctga ctaagttatc ttttaataat tggtattcct 10560
gtttattgct tgaagaattg ccggtcctat ttactcgttt taggactggt tcagaattca 10620
tcgatgctca ctcaaaggtc ggtaatacgg ttatccacag aatcagggga taacgcagga 10680
aagaac 10686

<210> 307

<211> 10686
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15148

<400> 307
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg tggatctacg cacgctcggc 2460
tacagccagc agcaacagga gaagatcaaa ccgaaggttc gttcgacagt ggcgcagcac 2520
cacgaggcac tggtcggcca cgggtttaca cacgcgcaca tcgttgcgtt aagccaacac 2580
ccggcagcgt tagggaccgt cgctgtcaag tatcaggaca tgatcgcagc gttgccagag 2640
gcgacacacg aagcgatcgt tggcgtcggc aaacagtggt ccggcgcacg cgctctggag 2700
gccttgctca cggtggcggg agagttgaga ggtccaccgt tacagttgga cacaggccaa 2760
cttctcaaga ttgcaaaacg tggcggcgtg accgcagtgg aggcagtgca tgcatggcgc 2820
aatgcactga cgggtgcccc gctcaacttg acccccccagc aggtggtggc catcgccagc 2880

aataatggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2940
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag 3000
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc 3060
cagcaggtgg tggccatcgc cagcaatggc ggtggcaagc aggcgctgga gacggtccag 3120
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 3180
gccagcaata atggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3240
tgccaggccc acggcttgac cccccagcag gtggtggcca tcgccagcaa tggcggtggc 3300
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 3360
acccccccagc aggtggtggc catcgccagc aataatggtg gcaagcaggc gctggagacg 3420
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 3480
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 3540
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 3600
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 3660
ggcttgaccc cccagcaggt ggtggccatc gccagcaata atggtggcaa gcaggcgctg 3720
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 3780
gtggtggcca tcgccagcaa taatggtggc aagcaggcgc tggagacggt ccagcggctg 3840
ttgccggtgc tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc 3900
aatggcggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 3960
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag 4020
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 4080
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 4140
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc ctcagcaggt ggtggccatc 4200
gccagcaatg gcggcggcag gccggcgctg gagagcattg ttgcccagtt atctcgccct 4260
gatccgggat ccgacatcac caaatctaaa atctctgaaa agatgaaagg ccagggccca 4320
tccggaaata ccaaatataa cgaagagttc ctgctgtacc tggccggctt tgtggacggt 4380
gacggtagca tcgttgctca gattaaacca catcagtctt gtaagtttaa acatcagcta 4440
cgtttgacct ttaacgtgac tcaaaagacc cagcgccgtt ggtttctgga caaactagtg 4500
gatgaaattg gcgttggtca tgtatacgat tctggtagcg tttcctacta ccagttaagc 4560
gaaatcaagc cgctgcacaa cttcctgact caactgcagc cgtttctgga actgaaacag 4620
aaacaggcaa acctggttct gaaaattatc gaacagctgc cgtctgcaaa agaatccccg 4680
gacaaattcc tggaagtttg tacctgggtg gatcaggttg cagctctgaa cgattctaag 4740
acgcgtaaaa ccacttctga aacctgtcgt gctgtgctgg acagcctgag cgagaagaag 4800
aaatcctccc cggcggccga ctgataactc gagcgatcct ctagacgagc tcctcgagcc 4860
tgcagcagct gaagctttgg acttcttcgc cagaggtttg gtcaagtctc caatcaaggt 4920
tgtcggcttg tctaccttgc cagaaattta cgaaaagatg gaaaagggtc aaatcgttgg 4980
tagatacgtt gttgacactt ctaaataagc gaatttctta tgatttatga tttttattat 5040
taaataagtt ataaaaaaaa taagtgtata caaattttaa agtgactctt aggttttaaa 5100
acgaaaattc ttattcttga gtaactcttt cctgtaggtc aggttgcttt ctcaggtata 5160
gcatgaggtc gctcttattg accacacctc taccggcatg caagcttggc gtaatcatgg 5220
tcatagctgt ttcctgtgtg aaattgttat ccgctcacaa ttccacacaa catacgagcc 5280
ggaagcataa agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaattgcg 5340
ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt gccagcagat ctattacatt 5400
atgggtggta tgttggaata aaaatcaact atcatctact aactagtatt tacgttacta 5460
gtatattatc atatacggtg ttagaagatg acgcaaatga tgagaaatag tcatctaaat 5520
tagtggaagc tgaaacgcaa ggattgataa tgtaatagga tcaatgaata ttaacatata 5580
aaatgatgat aataatattt atagaattgt gtagaattgc agattccctt ttatggattc 5640
ctaaatcctc gaggagaact tctagtatat ctacatacct aatattattg ccttattaaa 5700
aatggaatcc caacaattac atcaaaatcc acattctctt caaaatcaat tgtcctgtac 5760
ttccttgttc atgtgtgttc aaaaacgtta tatttatagg ataattatac tctatttctc 5820
aacaagtaat tggttgtttg gccgagcggt ctaaggcgcc tgattcaaga aatatcttga 5880
ccgcagttaa ctgtgggaat actcaggtat cgtaagatgc aagagttcga atctcttagc 5940
aaccattatt tttttcctca acataacgag aacacacagg ggcgctatcg cacagaatca 6000
aattcgatga ctggaaattt tttgttaatt tcagaggtcg cctgacgcat atacctttt 6060
caactgaaaa attgggagaa aaaggaaagg tgagagccgc ggaaccggct tttcatatag 6120
aatagagaag cgttcatgac taaatgcttg catcacaata cttgaagttg acaatattat 6180
ttaaggacct attgtttttt ccaataggtg gttagcaatc gtcttacttt ctaacttttc 6240
ttacctttta catttcagca atatatatat atatatttca aggatatacc attctaatgt 6300
ctgcccctaa gaagatcgtc gttttgccag gtgaccacgt tggtcaagaa atcacagccg 6360
aagccattaa ggttcttaaa gctatttctg atgttcgttc caatgtcaag ttcgatttcg 6420
aaaatcattt aattggtggt gctgctatcg atgctacagg tgtcccactt ccagatgagg 6480
cgctggaagc ctccaagaag gttgatgccg ttttgttagg tgctgtgggt ggtcctaaat 6540
ggggtaccgg tagtgttaga cctgaacaag gtttactaaa aatccgtaaa gaacttcaat 6600
tgtacgccaa cttaagacca tgtaactttg catccgactc tcttttagac ttatctccaa 6660
tcaagccaca atttgctaaa ggtactgact tcgttgttgt cagagaatta gtgggaggta 6720
tttactttgg taagagaaag gaagacgatg gtgatggtgt cgcttgggat agtgaacaat 6780
acaccgttcc agaagtgcaa agaatcacaa gaatggccgc tttcatggcc ctacaacatg 6840
agccaccatt gcctatttgg tccttggata aagctaatgt tttggcctct tcaagattat 6900
ggagaaaaac tgtggaggaa accatcaaga acgaattccc tacattgaag gttcaacatc 6960

aattgattga ttctgccgcc atgatcctag ttaagaaccc aacccaccta aatggtatta 7020
taatcaccag caacatgttt ggtgatatca tctccgatga agcctccgtt atcccaggtt 7080
ccttgggttt gttgccatct gcgtccttgg cctctttgcc agacaagaac accgcatttg 7140
gtttgtacga accatgccac ggttctgctc cagatttgcc aaagaataag gtcaacccta 7200
tcgccactat cttgtctgct gcaatgatgt tgaaattgtc attgaacttg cctgaagaag 7260
gtaaggccat tgaagatgca gttaaaaagg ttttggatgc aggtatcaga actggtgatt 7320
taggtggttc caacagtacc acggaagtcg gtgatgctgt cgccgaagaa gttaagaaaa 7380
tccttgctta aaaagattct ctttttttat gatatttgta cataaacttt ataaatgaaa 7440
ttcataatag aaacgacacg aaattacaaa atggaatatg ttcatagggt agacgaaact 7500
atatacgcaa tctacataca tttatcaaga aggagaaaaa ggaggatgta aaggaataca 7560
ggtaagcaaa ttgatactaa tggctcaacg tgataaggaa aaagaattgc actttaacat 7620
taatattgac aaggaggagg gcaccacaca aaaagttagg tgtaacagaa aatcatgaaa 7680
ctatgattcc taatttatat attggaggat tttctctaaa aaaaaaaaaa tacaacaaat 7740
aaaaaacact caatgacctg accatttgat ggagtttaag tcaatacctt cttgaaccat 7800
ttcccataat ggtgaaagtt ccctcaagaa ttttactctg tcagaaacgg ccttaacgac 7860
gtagtcgacc tcctcttcag tactaaatct accaatacca aatctgatgg aagaatgggc 7920
taatgcatca tccttaccca gcgcatgtaa aacataagaa ggttctaggg aagcagatgt 7980
acaggctgaa cccgaggata atgcgatatc ccttagtgcc atcaataaag attctccttc 8040
cacgtaggcg aaagaaacgt taacacaccc tggataacga tgatctggag atccgttcaa 8100
cgtggtatgt tcagcggata atagaccttt gactaattta tcggatagtc ttttgatgtg 8160
agcttggtcg ttgtcaaatt ctttcttcat caatctcgca gcttcaccaa atcccgctac 8220
caatgggggg gccaaagtac cagatctgct gcattaatga atcggccaac gcgcggggag 8280
aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt 8340
cgttcggctg cggcgagcgg tatcagcatc gatgaattcc acggactata gactatacta 8400
gtatactccg tctactgtac gatacacttc cgctcaggtc cttgtccttt aacgaggcct 8460
taccactctt ttgttactct attgatccag ctcagcaaag gcagtgtgat ctaagattct 8520
atcttcgcga tgtagtaaaa ctagctagac cgagaaagag actagaaatg caaaaggcac 8580
ttctacaatg gctgccatca ttattatccg atgtgacgct gcagcttctc aatgatattc 8640
gaatacgctt tgaggagata cagcctaata tccgacaaac tgttttacag atttacgatc 8700
gtacttgtta cccatcattg aattttgaac atccgaacct gggagttttc cctgaaacag 8760
atagtatatt tgaacctgta taataatata tagtctagcg ctttacggaa gacaatgtat 8820
gtatttcggt tcctggagaa actattgcat ctattgcata ggtaatcttg cacgtcgcat 8880
ccccggttca ttttctgcgt ttccatcttg cacttcaata gcatatcttt gttaacgaag 8940
catctgtgct tcattttgta gaacaaaaat gcaacgcgag agcgctaatt tttcaaacaa 9000
agaatctgag ctgcattttt acagaacaga aatgcaacgc gaaagcgcta ttttaccaac 9060
gaagaatctg tgcttcattt ttgtaaaaca aaaatgcaac gcgagagcgc taatttttca 9120
aacaaagaat ctgagctgca tttttacaga acagaaatgc aacgcgagag cgctatttta 9180
ccaacaaaga atctatactt ctttttttgtt ctacaaaaat gcatcccgag agcgctattt 9240
ttctaacaaa gcatcttaga ttacttttttt tctcctttgt gcgctctata atgcagtctc 9300
ttgataactt tttgcactgt aggtccgtta aggttagaag aaggctactt tggtgtctat 9360
tttctcttcc ataaaaaaag cctgactcca cttcccgcgt ttactgatta ctagcgaagc 9420
tgcgggtgca ttttttcaag ataaaggcat ccccgattat attctatacc gatgtggatt 9480
gcgcatactt tgtgaacaga aagtgatagc gttgatgatt cttcattggt cagaaaatta 9540
tgaacggttt cttctatttt gtctctatat actacgtata ggaaatgttt acattttcgt 9600
attgttttcg attcactcta tgaatagttc ttactacaat ttttttgtct aaagagtaat 9660
actagagata aacataaaaa atgtagaggt cgagtttaga tgcaagttca aggagcgaaa 9720
ggtggatggg taggttatat agggatatag cacagagata tatagcaaag agatactttt 9780
gagcaatgtt tgtggaagcg gtattcgcaa tattttagta gctcgttaca gtccggtgcg 9840
tttttggttt tttgaaagtg cgtcttcaga gcgcttttgg ttttcaaaag cgctctgaag 9900
ttcctatact ttctagagaa taggaacttc ggaataggaa cttcaaagcg tttccgaaaa 9960
cgagcgcttc cgaaaatgca acgcgagctg cgcacataca gctcactgtt cacgtcgcac 10020
ctatatctgc gtgttgcctg tatatatata tacatgagaa gaacggcata gtgcgtgttt 10080
atgcttaaat gcgtacttat atgcgtctat ttatgtagga tgaaaggtag tctagtacct 10140
cctgtgatat tatcccattc catgcggggt atcgtatgct tccttcagca ctacccttta 10200
gctgttctat atgctgccac tcctcaattg gattagtctc atccttcaat gctatcattt 10260
cctttgatat tggatcatat gcatagtacc gagaaactag tgcgaagtag tgatcaggta 10320
ttgctgttat ctgatgagta tacgttgtcc tggccacggc agaagcacgc ttatcgctcc 10380
aatttcccac aacattagtc aactccgtta ggcccttcat tgaaagaaat gaggtcatca 10440
aatgtcttcc aatgtgagat tttgggccat tttttatagc aaagattgaa taaggcgcat 10500
ttttcttcaa agctttattg tacgatctga ctaagttatc ttttaataat tggtattcct 10560
gtttattgct tgaagaattg ccggtcctat ttactcgttt taggactggt tcagaattca 10620
tcgatgctca ctcaaaggtc ggtaatacgg ttatccacag aatcagggga taacgcagga 10680
aagaac 10686

<210> 308

<211> 10988

<212> DNA
<213> Artificial sequence

<220>
<223> pCLS14894

<400> 308
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtttt 600
tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc 660
ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg 720
agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca 780
atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca 840
cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag 900
ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac 960
ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc 1020
agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct 1080
agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc 1140
gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg 1200
cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc 1260
gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat 1320
tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag 1380
tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat 1440
aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg 1500
cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca 1560
cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga 1620
aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc 1680
ttccttttc aatattattg aagcatttat cagggttatt gtctcatgag cggatacata 1740
tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 1800
ccacctgacg tcgacggatc gggagatctc ccgatcccct atggtgcact ctcagtacaa 1860
tctgctctga tgccgcatag ttaagccagt atctgctccc tgcttgtgtg ttggaggtcg 1920
ctgagtagtg cgcgagcaaa atttaagcta caacaaggca aggcttgacc gacaattgca 1980
tgaagaatct gcttagggtt aggcgttttg cgctgcttcg cgatgtacgg gccagatata 2040
cgcgttgaca ttgattattg actagttatt aatagtaatc aattacgggg tcattagttc 2100
atagcccata tatggagttc cgcgttacat aacttacggt aaatggcccg cctggctgac 2160
cgcccaacga cccccgccca ttgacgtcaa taatgacgta tgttcccata gtaacgccaa 2220
tagggacttt ccattgacgt caatgggtgg agtatttacg gtaaactgcc cacttggcag 2280
tacatcaagt gtatcatatg ccaagtacgc cccctattga cgtcaatgac ggtaaatggc 2340
ccgcctggca ttatgcccag tacatgacct tatgggactt tcctacttgg cagtacatct 2400
acgtattagt catcgctatt accatggtga tgcggttttg gcagtacatc aatgggcgtg 2460
gatagcggtt tgactcacgg ggatttccaa gtctccaccc cattgacgtc aatgggagtt 2520
tgttttggca ccaaaatcaa cgggactttc caaaatgtcg taacaactcc gccccattga 2580
cgcaaatggg cggtaggcgt gtacggtggg aggtctatat aagcagagct ctctggctaa 2640
ctagagaacc cactgcttac tggcttatcg aaatgaattc gactcactgt tgggagaccc 2700
aagctggcta gttaagctat cacaagtttg tacaaaaaag caggctggcg cgcccaattc 2760
gccctatagt gagtcgtatt acgtcgcgct cactggccgt cgttttacaa cgtcgtgact 2820
gggaaaaccc tggcgttacc caacttaatc gccttgcagc acatcccccct ttcgccagct 2880
ggcgtaatag cgaagaggcc cgcaccgaaa cgcccttccc aacagttgcg cagcctgaat 2940
ggcgaatggg agcgccctgt agcggcgcat taagcgcggc gggtgtggtg gttacgcgca 3000
gcgtgaccgc tacacttgcc agcgccctag cgcccgctcc tttcgctttc ttcccttcct 3060
ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcgggggctc cctttagggt 3120
tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggt gatggttggc 3180
ctgtagtggg ccatagccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 3240
ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc 3300
ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 3360
acaaaaattt aacgcgaatt ttaacaaaat attaacgctt acaatttagg tggcactttt 3420
cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc aaatatgtat 3480
ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag gaagagtatg 3540
agtattcaac atttccgtgt cgcccttatt ccctttttg cggcattttg ccttcctgtt 3600
tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt gggtgcacga 3660

gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa 3720
gaacgttttc caatgatgag cacttttaaa gttctgctat gtggcgcggt attatcccgt 3780
attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt 3840
gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc 3900
agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac aacgatcgga 3960
ggaccgaagg agctaaccgc ttttttgcac aacatggggg atcatgtaac tcgccttgat 4020
cgttgggaac cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct 4080
gtagcaatgg caacaacgtt gcgcaaacta ttaactggcg aactacttac tctagcttcc 4140
cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg 4200
gcccttccgg ctggctggtt tattgctgat aaatctggag ccggtgagcg tggttctcgc 4260
ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg 4320
acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca 4380
ctgattaagc attggtaact gtcagaccaa gtttactcat atatacttta gattgattta 4440
aaacttcatt tttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc 4500
aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa 4560
ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca 4620
ccgctaccag cggtggtttg tttgccggat caagagctac caactctttt tccgaaggta 4680
actggcttca gcagagcgca gataccaaat actgttcttc tagtgtagcc gtagttaggc 4740
caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaat cctgttacca 4800
gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta 4860
ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttggag 4920
cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag cgccacgctt 4980
cccgaaggga gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc 5040
acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac 5100
ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac 5160
gccagcaacg cggccttttt acggttcctg gcctttgct ggccttttgc tcacatggtc 5220
tttcctgcgt tatcccctga ttctgtggat aaccgtatta ccgcctttga gtgagctgat 5280
accgctcgcc gcagccgaac gaccgagcgc agcgagtcag tgagcgagga agcggagagc 5340
gcccaatacg caaaccgcct ctccccgcgc gttggccgat tcattaatgc agctggcacg 5400
acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg agttagctca 5460
ctcattaggc accccaggct ttacacttta tgcttccggc tcgtatgttg tgtggaattg 5520
tgagcggata acaatttcac acaggaaaca gctatgacca tgattacgcc aagcgcgtca 5580
attaaccctc actaaaggga acaaaagctg ttaattaagc gcgccacacg caaacacaaa 5640
tacacagcgg ccttgccacc atggtggatc tacgcacgct cggctacagc cagcagcaac 5700
aggagaagat caaaccgaag gttcgttcga cagtggcgca gcaccacgag gcactggtcg 5760
gccacgggtt tacacacgcg cacatcgttg cgttaagcca acacccggca gcgttaggga 5820
ccgtcgctgt caagtatcag gacatgatcg cagcgttgcc agaggcgaca cacgaagcga 5880
tcgttggcgt cggcaaacag tggtccggcg cacgcgctct ggaggccttg ctcacggtgg 5940
cgggagagtt gagaggtcca ccgttacagt tggacacagg ccaacttctc aagattgcaa 6000
aacgtggcgg cgtgaccgca gtggaggcag tgcatgcatg gcgcaatgca ctgacgggtg 6060
ccccgctcaa cttgaccccc cagcaggtgg tggccatcgc cagcaataat ggtggcaagc 6120
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 6180
cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg gagacggtcc 6240
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag gtggtggcca 6300
tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 6360
tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc aataatggtg 6420
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 6480
tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag gcgctggaga 6540
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgacccccc cagcaggtgg 6600
tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag cggctgttgc 6660
cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc gccagccacg 6720
atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg tgccaggccc 6780
acggcttgac cccccagcag gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc 6840
tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg acccccccagc 6900
aggtggtggc catcgccagc aataatggtg gcaagcaggc gctggagacg gtccagcggc 6960
tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg gccatcgcca 7020
gcaataatgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc 7080
aggcccacgg cttgacccccc cagcaggtgg tggccatcgc cagcaatggc ggtggcaagc 7140
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 7200
cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg gagacggtcc 7260
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag gtggtggcca 7320
tcgccagcaa tattggtggc aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc 7380
tgtgccaggc ccacggcttg acccctcagc aggtggtggc catcgccagc aatggcggcg 7440
gcaggccggc gctggagagc attgttgccc agttatctcg ccctgatccg ggatccgaca 7500
tcaccaaatc taaaatctct gaaaagatga aaggccaggg cccatccgga aataccaaat 7560
ataacgaaga gttcctgctg tacctggccg gctttgtgga cggtgacggt agcatcgttg 7620
ctcagattaa accacatcag tcttgtaagt ttaaacatca gctacgtttg acctttaacg 7680
tgactcaaaa gacccagcgc cgttggtttc tggacaaact agtggatgaa attggcgttg 7740

gtcatgtata cgattctggt agcgtttcct actaccagtt aagcgaaatc aagccgctgc 7800
acaacttcct gactcaactg cagccgtttc tggaactgaa acagaaacag gcaaacctgg 7860
ttctgaaaat tatcgaacag ctgccgtctg caaaagaatc cccggacaaa ttcctggaag 7920
tttgtacctg ggtggatcag gttgcagctc tgaacgattc taagacgcgt aaaaccactt 7980
ctgaaaccgt tcgtgctgtg ctggacagcc tgagcgagaa gaagaaatcc tccccggcgg 8040
ccggtgattc ctctgtttct aattccgagc acattgctcc tctgtctctg ccttcctctc 8100
ctccatctgt tggttctaac aaaaaattcc tgctgtatct tgctggattt gtggattctg 8160
atggctccat cattgctcag ataaaaccaa atcaatctta caagttcaaa cactacctct 8220
ccttgacctt ttgtgtcact cagaagacac aaagaaggtg gttcttggac aaattggttg 8280
atcgtattgg tgtgggctat gtccaggaca gaggctctgt gtcagactac cgtctgtctg 8340
aaattaagcc tcttcgtaac tttctcaccc aactgcaacc cttcttgaag ctcaaacaga 8400
agcaagcaaa tctggttttg aaaatcatcg agcaactgcc atctgccaag gagtcccctg 8460
acaagtttct tgaagtgtgt acttgggtgg atcaggttgc tgccttgaat gactccaaga 8520
ccagaaaaac cacctctgag actgtgaggg cagttctgga tagcctctct gagaagaaaa 8580
agtcctctcc ttagtaactc gagcgctagc acccagcttt cttgtacaaa gtggtgatct 8640
agagggcccg cggttcgaag gtaagcctat ccctaaccct ctcctcggtc tcgattctac 8700
gcgtaccggt tagtaatgag tttaaacggg ggaggctaac tgaaacacgg aaggagacaa 8760
taccggaagg aacccgcgct atgacggcaa taaaaagaca gaataaaacg cacgggtgtt 8820
gggtcgtttg ttcataaacg cggggttcgg tcccagggct ggcactctgt cgatacccca 8880
ccgagacccc attggggcca atacgcccgc gtttcttcct tttccccacc ccaccccca 8940
agttcgggtg aaggcccagg gctcgcagcc aacgtcgggg cggcaggccc tgccatagca 9000
gatctgcgca gctggggctc taggggtat ccccacgcgc cctgtagcgg cgcattaagc 9060
gcggcgggtg tggtggttac gcgcagcgtg accgctacac ttgccagcgc cctagcgccc 9120
gctcctttcg ctttcttccc ttcctttctc gccacgttcg ccggctttcc ccgtcaagct 9180
ctaaatcggg gcatcccttt agggttccga tttagtgctt tacggcacct cgaccccaaa 9240
aaacttgatt agggtgatgg ttcacgtagt gggccatcgc cctgatagac ggtttttcgc 9300
cctttgacgt tggagtccac gttctttaat agtggactct tgttccaaac tggaacaaca 9360
ctcaacccta tctcggtcta ttcttttgat ttataaggga ttttggggat ttcggcctat 9420
tggttaaaaa atgagctgat ttaacaaaaa tttaacgcga attaattctg tggaatgtgt 9480
gtcagttagg gtgtggaaag tccccaggct ccccagcagg cagaagtatg caaagcatgc 9540
atctcaatta gtcagcaacc aggtgtggaa agtccccagg ctccccagca ggcagaagta 9600
tgcaaagcat gcatctcaat tagtcagcaa ccatagtccc gcccctaact ccgcccatcc 9660
cgcccctaac tccgcccagt tccgcccatt ctccgcccca tggctgacta attttttta 9720
tttatgcaga ggccgaggcc gcctctgcct ctgagctatt ccagaagtag tgaggaggct 9780
tttttggagg cctaggcttt tgcaaaaagc tcccgggagc ttgtatatcc attttcggat 9840
ctgatcagca cgtgttgaca attaatcatc ggcatagtat atcggcatag tataatacga 9900
caaggtgagg aactaaacca tggccaagcc tttgtctcaa gaagaatcca ccctcattga 9960
aagagcaacg gctacaatca acagcatccc catctctgaa gactacagcg tcgccagcgc 10020
agctctctct agcgacggcc gcatcttcac tggtgtcaat gtatatcatt ttactggggg 10080
accttgtgca gaactcgtgg tgctgggcac tgctgctgct gcggcagctg gcaacctgac 10140
ttgtatcgtc gcgatcggaa atgagaacag gggcatcttg agcccctgcg gacggtgccg 10200
acaggtgctt ctcgatctgc atcctgggat caaagccata gtgaaggaca gtgatggaca 10260
gccgacggca gttgggattc gtgaattgct gccctctggt tatgtgtggg agggctaagc 10320
acttcgtggc cgaggagcag gactgacacg tgctacgaga tttcgattcc accgccgcct 10380
tctatgaaag gttgggcttc ggaatcgttt tccgggacgc cggctggatg atcctccagc 10440
gcggggatct catgctggag ttcttcgccc accccaactt gtttattgca gcttataatg 10500
gttacaaata aagcaatagc atcacaaatt tcacaaataa agcatttttt tcactgcatt 10560
ctagttgtgg tttgtccaaa ctcatcaatg tatcttatca tgtctgtata ccgtcgacct 10620
ctagctagag cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc 10680
tcacaattcc acacaacata cgagccggaa gcataaagtg taaagcctgg ggtgcctaat 10740
gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 10800
tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg 10860
ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag 10920
cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag 10980
gaaagaac 10988

<210> 309

<211> 10988
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS14895

<400> 309
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60

ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtttt 600
tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc 660
ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg 720
agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca 780
atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca 840
cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag 900
ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac 960
ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc 1020
agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct 1080
agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc 1140
gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg 1200
cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc 1260
gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat 1320
tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag 1380
tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat 1440
aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg 1500
cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca 1560
cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga 1620
aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc 1680
ttccttttc aatattattg aagcatttat cagggttatt gtctcatgag cggatacata 1740
tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 1800
ccacctgacg tcgacggatc gggagatctc ccgatcccct atggtgcact ctcagtacaa 1860
tctgctctga tgccgcatag ttaagccagt atctgctccc tgcttgtgtg ttggaggtcg 1920
ctgagtagtg cgcgagcaaa atttaagcta caacaaggca aggcttgacc gacaattgca 1980
tgaagaatct gcttagggtt aggcgttttg cgctgcttcg cgatgtacgg gccagatata 2040
cgcgttgaca ttgattattg actagttatt aatagtaatc aattacgggg tcattagttc 2100
atagcccata tatggagttc cgcgttacat aacttacggt aaatggcccg cctggctgac 2160
cgcccaacga cccccgccca ttgacgtcaa taatgacgta tgttcccata gtaacgccaa 2220
tagggacttt ccattgacgt caatgggtgg agtatttacg gtaaactgcc cacttggcag 2280
tacatcaagt gtatcatatg ccaagtacgc cccctattga cgtcaatgac ggtaaatggc 2340
ccgcctggca ttatgcccag tacatgacct tatgggactt tcctacttgg cagtacatct 2400
acgtattagt catcgctatt accatggtga tgcggttttg gcagtacatc aatgggcgtg 2460
gatagcggtt tgactcacgg ggatttccaa gtctccaccc cattgacgtc aatgggagtt 2520
tgttttggca ccaaaatcaa cgggactttc caaaatgtcg taacaactcc gccccattga 2580
cgcaaatggg cggtaggcgt gtacggtggg aggtctatat aagcagagct ctctggctaa 2640
ctagagaacc cactgcttac tggcttatcg aaatgaattc gactcactgt tgggagaccc 2700
aagctggcta gttaagctat cacaagtttg tacaaaaaag caggctggcg cgcccaattc 2760
gccctatagt gagtcgtatt acgtcgcgct cactggccgt cgttttacaa cgtcgtgact 2820
gggaaaaccc tggcgttacc caacttaatc gccttgcagc acatccccct ttcgccagct 2880
ggcgtaatag cgaagaggcc cgcaccgaaa cgcccttccc aacagttgcg cagcctgaat 2940
ggcgaatggg agcgccctgt agcggcgcat taagcgcggc gggtgtggtg gttacgcgca 3000
gcgtgaccgc tacacttgcc agcgccctag cgcccgctcc tttcgctttc ttcccttcct 3060
ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcgggggctc cctttagggt 3120
tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggt gatggttggc 3180
ctgtagtggg ccatagccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 3240
ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc 3300
ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 3360
acaaaaattt aacgcgaatt ttaacaaaat attaacgctt acaatttagg tggcacttt 3420
cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc aaatatgtat 3480
ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag gaagagtatg 3540
agtattcaac atttccgtgt cgcccttatt cccttttttg cggcattttg ccttcctgtt 3600
tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt gggtgcacga 3660
gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa 3720
gaacgttttc caatgatgag cacttttaaa gttctgctat gtggcgcggt attatcccgt 3780
attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt 3840
gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc 3900
agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac aacgatcgga 3960
ggaccgaagg agctaaccgc ttttttgcac aacatggggg atcatgtaac tcgccttgat 4020
cgttgggaac cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct 4080
gtagcaatgg caacaacgtt gcgcaaacta ttaactggcg aactacttac tctagcttcc 4140

```
cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg   4200
gcccttccgg ctggctggtt tattgctgat aaatctggag ccggtgagcg tggttctcgc   4260
ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg   4320
acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca   4380
ctgattaagc attggtaact gtcagaccaa gtttactcat atatacttta gattgattta   4440
aaacttcatt tttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc   4500
aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa   4560
ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca   4620
ccgctaccag cggtggtttg tttgccggat caagagctac caactctttt tccgaaggta   4680
actggcttca gcagagcgca gataccaaat actgttcttc tagtgtagcc gtagttaggc   4740
caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaat cctgttacca   4800
gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta   4860
ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttggag   4920
cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag cgccacgctt   4980
cccgaaggga gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc   5040
acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac   5100
ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac   5160
gccagcaacg cggccttttt acggttcctg gccttttgct ggccttttgc tcacatggtc   5220
tttcctgcgt tatcccctga ttctgtggat aaccgtatta ccgcctttga gtgagctgat   5280
accgctcgcc gcagccgaac gaccgagcgc agcgagtcag tgagcgagga agcggagagc   5340
gcccaatacg caaaccgcct ctccccgcgc gttggccgat tcattaatgc agctggcacg   5400
acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg agttagctca   5460
ctcattaggc accccaggct ttacacttta tgcttccggc tcgtatgttg tgtggaattg   5520
tgagcggata acaatttcac acaggaaaca gctatgacca tgattacgcc aagcgcgtca   5580
attaaccctc actaaaggga acaaaagctg ttaattaagc gcgccacacg caaacacaaa   5640
tacacagcgg ccttgccacc atggtggatc tacgcacgct cggctacagc cagcagcaac   5700
aggagaagat caaaccgaag gttcgttcga cagtggcgca gcaccacgag gcactggtcg   5760
gccacgggtt tacacacgcg cacatcgttg cgttaagcca acacccggca gcgttaggga   5820
ccgtcgctgt caagtatcag gacatgatcg cagcgttgcc agaggcgaca cacgaagcga   5880
tcgttggcgt cggcaaacag tggtccggcg cacgcgctct ggaggccttg ctcacggtgg   5940
cgggagagtt gagaggtcca ccgttacagt tggacacagg ccaacttctc aagattgcaa   6000
aacgtggcgg cgtgaccgca gtggaggcag tgcatgcatg gcgcaatgca ctgacgggtg   6060
ccccgctcaa cttgaccccc cagcaggtgg tggccatcgc cagcaataat ggtggcaagc   6120
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc   6180
cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc   6240
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag gtggtggcca   6300
tcgccagcaa tattggtggc aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc   6360
tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc aatggcggtg   6420
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct   6480
tgacccccca gcaggtggtg gccatcgcca gcaataatgg tggcaagcag gcgctggaga   6540
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg   6600
tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag cggctgttgc   6660
cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc gccagcaatg   6720
gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg tgccaggccc   6780
acggcttgac cccccagcag gtggtggcca tcgccagcaa taatggtggc aagcaggcgc   6840
tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg acccccccagc   6900
aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg gtccagcggc   6960
tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg gccatcgcca   7020
gcaataatgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc   7080
aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat ggcggcaagc   7140
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc   7200
cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc   7260
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag gtggtggcca   7320
tcgccagcaa taatggtggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc   7380
tgtgccaggc ccacggcttg acccctcagc aggtggtggc catcgccagc aatggcggcg   7440
gcaggccggc gctggagagc attgttgccc agttatctcg ccctgatccg ggatccgaca   7500
tcaccaaatc taaaatctct gaaaagatga aaggccaggg cccatccgga aataccaaat   7560
ataacgaaga gttcctgctg tacctggccg gctttgtgga cggtgacggt agcatcgttg   7620
ctcagattaa accacatcag tcttgtaagt ttaaacatca gctacgtttg acctttaacg   7680
tgactcaaaa gacccagcgc cgttggtttc tggacaaact agtggatgaa attggcgttg   7740
gtcatgtata cgattctggt agcgtttcct actaccagtt aagcgaaatc aagccgctgc   7800
acaacttcct gactcaactg cagccgtttc tggaactgaa acagaaacag gcaaacctgg   7860
ttctgaaaat tatcgaacag ctgccgtctg caaaagaatc cccggacaaa ttcctggaag   7920
tttgtacctg ggtggatcag gttgcagctc tgaacgattc taagacgcgt aaaaccactt   7980
ctgaaaccgt tcgtgctgtg ctggacagcc tgagcgagaa gaagaaatcc tccccggcgg   8040
ccggtgattc ctctgtttct aattccgagc acattgctcc tctgtctctg ccttcctctc   8100
ctccatctgt tggttctaac aaaaaattcc tgctgtatct tgctggattt gtggattctg   8160
atggctccat cattgctcag ataaaaccaa atcaatctta caagttcaaa cactacctct   8220
```

```
ccttgacctt ttgtgtcact cagaagacac aaagaaggtg gttcttggac aaattggttg   8280
atcgtattgg tgtgggctat gtccaggaca gaggctctgt gtcagactac cgtctgtctg   8340
aaattaagcc tcttcgtaac tttctcaccc aactgcaacc cttcttgaag ctcaaacaga   8400
agcaagcaaa tctggttttg aaaatcatcg agcaactgcc atctgccaag gagtcccctg   8460
acaagtttct tgaagtgtgt acttgggtgg atcaggttgc tgccttgaat gactccaaga   8520
ccagaaaaac cacctctgag actgtgaggg cagttctgga tagcctctct gagaagaaaa   8580
agtcctctcc ttagtaactc gagcgctagc acccagcttt cttgtacaaa gtggtgatct   8640
agagggcccg cggttcgaag gtaagcctat ccctaaccct ctcctcggtc tcgattctac   8700
gcgtaccggt tagtaatgag tttaaacggg ggaggctaac tgaaacacgg aaggagacaa   8760
taccggaagg aacccgcgct atgacggcaa taaaaagaca gaataaaacg cacgggtgtt   8820
gggtcgtttg ttcataaacg cggggttcgg tcccagggct ggcactctgt cgatacccca   8880
ccgagacccc attggggcca atacgcccgc gtttcttcct tttccccacc ccacccccca   8940
agttcgggtg aaggcccagg gctcgcagcc aacgtcgggg cggcaggccc tgccatagca   9000
gatctgcgca gctggggctc taggggtat ccccacgcgc cctgtagcgg cgcattaagc   9060
gcggcgggtg tggtggttac gcgcagcgtg accgctacac ttgccagcgc cctagcgccc   9120
gctcctttcg ctttcttccc ttcctttctc gccacgttcg ccggctttcc ccgtcaagct   9180
ctaaatcggg gcatcccttt agggttccga tttagtgctt tacggcacct cgaccccaaa   9240
aaacttgatt agggtgatgg ttcacgtagt gggccatcgc cctgatagac ggtttttcgc   9300
cctttgacgt tggagtccac gttctttaat agtggactct tgttccaaac tggaacaaca   9360
ctcaacccta tctcggtcta ttcttttgat ttataaggga ttttggggat ttcggcctat   9420
tggttaaaaa atgagctgat ttaacaaaaa tttaacgcga attaattctg tggaatgtgt   9480
gtcagttagg gtgtggaaag tccccaggct ccccagcagg cagaagtatg caaagcatgc   9540
atctcaatta gtcagcaacc aggtgtggaa agtccccagg ctccccagca ggcagaagta   9600
tgcaaagcat gcatctcaat tagtcagcaa ccatagtccc gcccctaact ccgcccatcc   9660
cgcccctaac tccgcccagt tccgcccatt ctccgcccca tggctgacta attttttta    9720
tttatgcaga ggccgaggcc gcctctgcct ctgagctatt ccagaagtag tgaggaggct   9780
tttttggagg cctaggcttt tgcaaaaagc tcccgggagc ttgtatatcc attttcggat   9840
ctgatcagca cgtgttgaca attaatcatc ggcatagtat atcggcatag tataatacga   9900
caaggtgagg aactaaacca tggccaagcc tttgtctcaa gaagaatcca ccctcattga   9960
aagagcaacg gctacaatca acagcatccc catctctgaa gactacagcg tcgccagcgc  10020
agctctctct agcgacggcc gcatcttcac tggtgtcaat gtatatcatt ttactggggg  10080
accttgtgca gaactcgtgg tgctgggcac tgctgctgct gcggcagctg gcaacctgac  10140
ttgtatcgtc gcgatcggaa atgagaacag gggcatcttg agcccctgcg gacggtgccg  10200
acaggtgctt ctcgatctgc atcctgggat caaagccata gtgaaggaca gtgatggaca  10260
gccgacggca gttgggattc gtgaattgct gccctctggt tatgtgtggg agggctaagc  10320
acttcgtggc cgaggagcag gactgacacg tgctacgaga tttcgattcc accgccgcct  10380
tctatgaaag gttgggcttc ggaatcgttt tccgggacgc cggctggatg atcctccagc  10440
gcggggatct catgctggag ttcttcgccc accccaactt gtttattgca gcttataatg  10500
gttacaaata aagcaatagc atcacaaatt tcacaaataa agcatttttt tcactgcatt  10560
ctagttgtgg tttgtccaaa ctcatcaatg tatcttatca tgtctgtata ccgtcgacct  10620
ctagctagag cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc  10680
tcacaattcc acacaacata cgagccggaa gcataaagtg taaagcctgg ggtgcctaat  10740
gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc  10800
tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg  10860
ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag  10920
cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag  10980
gaaagaac                                                           10988
```

<210> 310

<211> 6626
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS8982

<400> 310

```
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt     60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg    120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc    180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc    240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc    300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac    360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt    420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct    480
aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc    540
```

```
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtttt    600
tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc    660
ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg    720
agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca    780
atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca    840
cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag    900
ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac    960
ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc   1020
agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct   1080
agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc   1140
gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg   1200
cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc   1260
gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat   1320
tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag   1380
tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat   1440
aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg   1500
cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca   1560
cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga   1620
aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc   1680
ttccttttttc aatattattg aagcatttat cagggttatt gtctcatgag cggatacata   1740
tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg   1800
ccacctgacg tcgacggatc gggagatctc ccgatcccct atggtgcact ctcagtacaa   1860
tctgctctga tgccgcatag ttaagccagt atctgctccc tgcttgtgtg ttggaggtcg   1920
ctgagtagtg cgcgagcaaa atttaagcta caacaaggca aggcttgacc gacaattgca   1980
tgaagaatct gcttagggtt aggcgttttg cgctgcttcg cgatgtacgg gccagatata   2040
cgcgttgaca ttgattattg actagttatt aatagtaatc aattacgggg tcattagttc   2100
atagcccata tatggagttc cgcgttacat aacttacggt aaatggcccg cctggctgac   2160
cgcccaacga cccccgccca ttgacgtcaa taatgacgta tgttcccata gtaacgccaa   2220
tagggacttt ccattgacgt caatgggtgg agtatttacg gtaaactgcc cacttggcag   2280
tacatcaagt gtatcatatg ccaagtacgc cccctattga cgtcaatgac ggtaaatggc   2340
ccgcctggca ttatgcccag tacatgacct tatgggactt tcctacttgg cagtacatct   2400
acgtattagt catcgctatt accatggtga tgcggttttg gcagtacatc aatgggcgtg   2460
gatagcggtt tgactcacgg ggatttccaa gtctccaccc cattgacgtc aatgggagtt   2520
tgttttggca ccaaaatcaa cgggactttc caaaatgtcg taacaactcc gccccattga   2580
cgcaaatggg cggtaggcgt gtacggtggg aggtctatat aagcagagct ctctggctaa   2640
ctagagaacc cactgcttac tggcttatcg aaatgaattc gactcactgt tgggagaccc   2700
aagctggcta gttaagctat cacaagtttg tacaaaaaag caggctggcg cgcctacaca   2760
gcggccttgc caccatgggt tccgaggcac cccgggccga gacctttgtc ttcctggacc   2820
tggaagccac tgggctcccc agtgtggagc ccgagattgc cgagctgtcc ctctttgctg   2880
tccaccgctc ctccctggag aacccggagc acgacgagtc tggtgcccta gtattgcccc   2940
gggtcctgga caagctcacg ctgtgcatgt gcccggagcg ccccttcact gccaaggcca   3000
gcgagatcac cggcctgagc agtgagggcc tggcgcgatg ccggaaggct ggctttgatg   3060
gcgccgtggt gcggacgctg caggccttcc tgagccgcca ggcagggccc atctgccttg   3120
tggcccacaa tggctttgat tatgatttcc ccctgctgtg tgccgagctg cggcgcctgg   3180
gtgcccgcct gccccgggac actgtctgcc tggacacgct gccggccctg cggggcctgg   3240
accgcgccca cagccacggc acccgggccc ggggccgcca gggttacagc ctcggcagcc   3300
tcttccaccg ctacttccgg gcagagccaa gcgcagccca ctcagccgag ggcgacgtgc   3360
acaccctgct cctgatcttc ctgcaccgcg ccgcagagct gctcgcctgg gccgatgagc   3420
aggcccgtgg gtgggcccac atcgagccca tgtacttgcc gcctgatgac cccagcctgg   3480
aggcgactcc tccacagacc ggtctggatg ttccttactc cgaggcaccc cgggccgaga   3540
cctttgtctt cctggacctg gaagccactg ggctccccag tgtggagccc gagattgccg   3600
agctgtccct ctttgctgtc caccgctcct ccctggagaa cccggagcac gacgagtctg   3660
gtgccctagt attgccccgg gtcctggaca agctcacgct gtgcatgtgc ccggagcgcc   3720
ccttcactgc caaggccagc gagatcaccg gcctgagcag tgagggcctg gcgcgatgcc   3780
ggaaggctgg ctttgatggc gccgtggtgc ggacgctgca ggccttcctg agccgccagg   3840
cagggcccat ctgccttgtg gcccacaatg gctttgatta tgatttcccc ctgctgtgtg   3900
ccgagctgcg gcgcctgggt gcccgcctgc cccgggacac tgtctgcctg gacacgctgc   3960
cggccctgcg gggcctggac cgcgcccaca gccacggcac ccgggcccgg ggccgccagg   4020
gttacagcct cggcagcctc ttccaccgct acttccgggc agagccaagc gcagcccact   4080
cagccgaggg cgacgtgcac accctgctcc tgatcttcct gcaccgcgcc gcagagctgc   4140
tcgcctgggc cgatgagcag gcccgtgggt gggcccacat cgagcccatg tacttgccgc   4200
ctgatgaccc cagcctggag gcggccgact gataactcga gcgctagcac ccagctttct   4260
tgtacaaagt ggtgatctag agggcccgcg gttcgaaggt aagcctatcc ctaaccctct   4320
cctcggtctc gattctacgc gtaccggtta gtaatgagtt taaacggggg aggctaactg   4380
aaacacggaa ggagacaata ccggaaggaa cccgcgctat gacggcaata aaaagacaga   4440
ataaaacgca cgggtgttgg gtcgtttgtt cataaacgcg gggttcggtc ccagggctgg   4500
cactctgtcg ataccccacc gagaccccat tggggccaat acgcccgcgt ttcttccttt   4560
tccccacccc acccccccaag ttcgggtgaa ggcccagggc tcgcagccaa cgtcggggcg   4620
```

gcaggccctg ccatagcaga tctgcgcagc tggggctcta gggggtatcc ccacgcgccc 4680
tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt 4740
gccagcgccc tagcgcccgc tcctttcgct ttcttccctt cctttctcgc cacgttcgcc 4800
ggctttcccc gtcaagctct aaatcggggc atcccttag ggttccgatt tagtgcttta 4860
cggcacctcg accccaaaaa acttgattag ggtgatggtt cacgtagtgg gccatcgccc 4920
tgatagacgg tttttcgccc tttgacgttg gagtccacgt tctttaatag tggactcttg 4980
ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt ataagggatt 5040
ttggggattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt taacgcgaat 5100
taattctgtg gaatgtgtgt cagttagggt gtggaaagtc cccaggctcc ccagcaggca 5160
gaagtatgca aagcatgcat ctcaattagt cagcaaccag gtgtggaaag tccccaggct 5220
ccccagcagg cagaagtatg caaagcatgc atctcaatta gtcagcaacc atagtcccgc 5280
ccctaactcc gcccatcccg cccctaactc cgcccagttc cgcccattct ccgccccatg 5340
gctgactaat ttttttttatt tatgcagagg ccgaggccgc ctctgcctct gagctattcc 5400
agaagtagtg aggaggcttt tttggaggcc taggcttttg caaaaagctc ccgggagctt 5460
gtatatccat tttcggatct gatcagcacg tgttgacaat taatcatcgg catagtatat 5520
cggcatagta taatacgaca aggtgaggaa ctaaaccatg gccaagcctt tgtctcaaga 5580
agaatccacc ctcattgaaa gagcaacggc tacaatcaac agcatcccca tctctgaaga 5640
ctacagcgtc gccagcgcag ctctctctag cgacggccgc atcttcactg gtgtcaatgt 5700
atatcatttt actgggggac cttgtgcaga actcgtggtg ctgggcactg ctgctgctgc 5760
ggcagctggc aacctgactt gtatcgtcgc gatcggaaat gagaacaggg gcatcttgag 5820
cccctgcgga cggtgccgac aggtgcttct cgatctgcat cctgggatca aagccatagt 5880
gaaggacagt gatggacagc cgacggcagt tgggattcgt gaattgctgc cctctggtta 5940
tgtgtgggag ggctaagcac ttcgtggccg aggagcagga ctgacacgtg ctacgagatt 6000
tcgattccac cgccgccttc tatgaaaggt tgggcttcgg aatcgttttc cgggacgccg 6060
gctggatgat cctccagcgc ggggatctca tgctggagtt cttcgcccac cccaacttgt 6120
ttattgcagc ttataatggt tacaaataaa gcaatagcat cacaaatttc acaaataaag 6180
catttttttc actgcattct agttgtggtt tgtccaaact catcaatgta tcttatcatg 6240
tctgtatacc gtcgacctct agctagagct tggcgtaatc atggtcatag ctgtttcctg 6300
tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgta 6360
aagcctgggg tgcctaatga gtgagctaac tcacattaat tgcgttgcgc tcactgcccg 6420
ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga 6480
gaggcggttt gcgtattggg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg 6540
tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag 6600
aatcagggga taacgcagga aagaac 6626

<210> 311

<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> TRBC2F3 oligonucleotide

<400> 311
gtgggagaga ccagagctac ctg 23

<210> 312

<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> TRBC2R3B oligonucleotide

<400> 312
ccttactatt tgtgtggtca atgg 24

<210> 313

<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> TRBC2F4 oligonucleotide

<400> 313
aggtacccac ctgtccttcc tc 22

<210> 314

<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> TRBC2R4B oligonucleotide

<400> 314
atggatgttg tgatgagcct tac 23

<210> 315

<211> 41
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs5

<400> 315
tgtttatggt tacttatagg tatatataaa cctaaccctc t 41

<210> 316

<211> 42
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs6

<400> 316
tgtttatggt tacttataag gtatatataa acctaaccct ct 42

<210> 317

<211> 43
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs7

<400> 317
tgtttatggt tacttataag gtactatata aacctaaccc tct 43

<210> 318

<211> 44
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs8

<400> 318
tgtttatggt tacttatgaa ggtactatat aaacctaacc ctct 44

<210> 319

<211> 45
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs9

<400> 319
tgtttatggt tacttatgaa ggtacctata taaacctaac cctct 45

<210> 320

<211> 46
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs10

<400> 320
tgtttatggt tacttattga aggtacctat ataaacctaa ccctct 46

<210> 321

<211> 47
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs11

<400> 321
tgtttatggt tacttattga aggtacctta tataaaccta accctct 47

<210> 322

<211> 48
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs12

<400> 322
tgtttatggt tacttatatg aaggtacctt atataaacct aaccctct 48

<210> 323

<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs13

<400> 323
tgtttatggt tacttatatg aaggtacctt tatataaacc taaccctct 49

<210> 324

<211> 50
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs14

<400> 324
tgtttatggt tacttatcat gaaggtacct ttatataaac ctaaccctct 50

<210> 325

<211> 51
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs15

<400> 325
tgtttatggt tacttattag catgaaggta cctatataaa cctaaccctc t 51

<210> 326

<211> 52
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs16

<400> 326
tgtttatggt tacttatgca tgaaggtacc ttgtatataa acctaaccct ct 52

<210> 327

<211> 53
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs17

<400> 327
tgtttatggt tacttatgca tgaaggtacc ttgttatata aacctaaccc tct 53

<210> 328

<211> 54
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs18

<400> 328
tgtttatggt tacttatagc atgaaggtac cttgttatat aaacctaacc ctct 54

<210> 329

<211> 55
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs19

<400> 329
tgtttatggt tacttatagc atgaaggtac cttgtctata taaacctaac cctct 55

<210> 330

<211> 56
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs20

<400> 330
tgtttatggt tacttattag catgaaggta ccttgtctat ataaacctaa ccctct 56

<210> 331

<211> 57
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs21

<400> 331
tgtttatggt tacttattag catgaaggta ccttgtcgta tataaaccta accctct 57

<210> 332

<211> 58
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs22

<400> 332
tgtttatggt tacttattag catgaaggta ccttgtcgtt atataaacct aaccctct 58

<210> 333

<211> 59
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs23

<400> 333
tgtttatggt tacttatcta gcatgaaggt accttgtcgt tatataaacc taaccctct 59

<210> 334

<211> 60
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs24

<400> 334
tgtttatggt tacttatcta gcatgaaggt accttgtcgt ttatataaac ctaaccctct 60

<210> 335

<211> 61
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs25

<400> 335
tgtttatggt tacttatact agcatgaagg taccttgtcg tttatataaa cctaaccctc 60
t 61

<210> 336

<211> 62
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs26

<400> 336
tgtttatggt tacttatact agcatgaagg taccttgtcg ttgtatataa acctaaccct 60
ct 62

<210> 337

<211> 63
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs27

<400> 337
tgtttatggt tacttatcac tagcatgaag gtaccttgtc gttgtatata aacctaaccc 60
tct 63

<210> 338

<211> 64
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs28

<400> 338
tgtttatggt tacttatcac tagcatgaag gtaccttgtc gttgatatat aaacctaacc 60
ctct 64

<210> 339

<211> 65
<212> DNA

<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs29

<400> 339
tgtttatggt tacttatcca ctagcatgaa ggtaccttgt cgttgatata taaacctaac 60
cctct 65

<210> 340

<211> 66
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs30

<400> 340
tgtttatggt tacttatcca ctagcatgaa ggtaccttgt cgttgattat ataaacctaa 60
ccctct 66

<210> 341

<211> 67
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs31

<400> 341
tgtttatggt tacttatacc actagcatga aggtaccttg tcgttgatta tataaaccta 60
accctct 67

<210> 342

<211> 68
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs32

<400> 342
tgtttatggt tacttatacc actagcatga aggtaccttg tcgttgattt atataaacct 60
aaccctct 68

<210> 343

<211> 69
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs33

<400> 343
tgtttatggt tacttatgac cactagcatg aaggtacctt gtcgttgatt tatataaacc 60
taaccctct 69

<210> 344

<211> 70
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs34

<400> 344
tgtttatggt tacttatgac cactagcatg aaggtacctt gtcgttgatt ctatataaac 60
ctaaccctct 70

<210> 345

<211> 71
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs35

<400> 345
tgtttatggt tacttattga ccactagcat gaaggtacct tgtcgttgat tctatataaa 60
cctaaccctc t 71

<210> 346

<211> 72
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs36

<400> 346
tgtttatggt tacttattga ccactagcat gaaggtacct tgtcgttgat tcatatataa 60
acctaaccct ct 72

<210> 347

<211> 73
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs37

<400> 347
tgtttatggt tacttatctg accactagca tgaaggtacc ttgtcgttga ttcatatata 60
aacctaaccc tct 73

<210> 348

<211> 74
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs38

<400> 348
tgtttatggt tacttatctg accactagca tgaaggtacc ttgtcgttga ttcagtatat 60
aaacctaacc ctct 74

<210> 349

<211> 75
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs39

<400> 349
tgtttatggt tacttattct gaccactagc atgaaggtac cttgtcgttg attcagtata 60
taaacctaac cctct 75

<210> 350

<211> 76
<212> DNA
<213> Artificial sequence

<220>
<223> hybrid RagT2-R/AvrBs40

<400> 350
tgtttatggt tacttattct gaccactagc atgaaggtac cttgtcgttg attcagttat 60
ataaacctaa ccctct 76

<210> 351

<211> 5864
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15795

<400> 351
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggtg gatctacgca cgctcggcta cagccagcag 480
caacaggaga agatcaaacc gaaggttcgt tcgacagtgg cgcagcacca cgaggcactg 540
gtcggccacg ggtttacaca cgcgcacatc gttgcgttaa gccaacaccc ggcagcgtta 600
gggaccgtcg ctgtcaagta tcaggacatg atcgcagcgt tgccagaggc gacacacgaa 660
gcgatcgttg gcgtcggcaa acagtggtcc ggcgcacgcg ctctggaggc cttgctcacg 720
gtggcgggag agttgagagg tccaccgtta cagttggaca caggccaact tctcaagatt 780
gcaaaacgtg gcggcgtgac cgcagtggag gcagtgcatg catggcgcaa tgcactgacg 840
ggtgccccgc tcaacttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 900
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 960
acccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 1020
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 1080
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 1140
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 1200
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 1260
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg 1320
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag 1380
gtggtggcca tcgccagcaa tattggtggc aagcaggcgc tggagacggt gcaggcgctg 1440
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 1500
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 1560
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag 1620
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 1680
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 1740
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 1800

```
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg   1860
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc   1920
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg   1980
accccggagc aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg   2040
gtgcaggcgc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg   2100
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg   2160
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat   2220
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac   2280
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg   2340
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag   2400
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg   2460
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc   2520
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag   2580
gcccacggct tgacccctca gcaggtggtg gccatcgcca gcaatggcgg cggcaggccg   2640
gcgctggaga gcattgttgc ccagttatct cgccctgatc cgggatcctc cggaccgaac   2700
agaggcgtta ccaagcaact agttaaatcc gaactggaag agaaaaagtc cgagttgcgt   2760
cataaactga agtacgttcc acatgaatac attgaactga tcgagattgc ccgtaacagc   2820
acccaggaca gaatcctgga aatgaaagtg atggagtttt tcatgaaagt gtacggctac   2880
agaggcaaac atttgggcgg tagccgtaaa ccggatggtg ccatctacac cgttggcagc   2940
ccgatcgatt acggcgtgat cgttgatacc aaagcctact ccggtggcta caacctgccg   3000
attggccagg ctgacgaaat gcagagatac gtggaagaga accagactcg taacaagcat   3060
atcaacccga acgaatggtg gaaagtctac ccgagttccg tgactgagtt caaatttctg   3120
tttgtgagcg gccatttcaa aggcaactac aaagcccagc tgaccagact gaaccatatc   3180
accaactgca acggtgccgt cttgagcgtg gaagaactgt tgattggtgg cgaaatgatc   3240
aaggctggca ccctgactct ggaggaagtt agacgtaaat tcaacaacgg tgaaatcaac   3300
tttggatcct cggccgactg ataagtcgac caccaccacc accaccactg ataagagctc   3360
ctcgagcctg cagcagctga agctttggac ttctggcgcg cccaattcgc cctatagtga   3420
gtcgtattac gtcgcgctca ctggccgtcg ttttacaacg tcgtgactgg gaaaaccctg   3480
gcgttaccca acttaatcgc cttgcagcac atcccccttt cgccagctgg cgtaatagcg   3540
aagaggcccg caccgaaacg cccttcccaa cagttgcgca gcctgaatgg cgaatgggag   3600
cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta   3660
cacttgccag cgccctagcg cccgctcctt tcgctttctt cccttccttt ctcgccacgt   3720
tcgccggctt tccccgtcaa gctctaaatc gggggctccc tttagggttc cgatttagtg   3780
ctttacggca cctcgacccc aaaaaacttg attagggtga tggttggcct gtagtgggcc   3840
atagccctga tagacggttt ttcgcccttt gacgttggag tccacgttct ttaatagtgg   3900
actcttgttc caaactggaa caacactcaa ccctatctcg gtctattctt ttgatttata   3960
agggattttg ccgatttcgg cctattggtt aaaaaatgag ctgatttaac aaaaatttaa   4020
cgcgaatttt aacaaaatat taacgcttac aatttaggtg gcacttttcg gggaaatgtg   4080
cgcggaaccc ctatttgttt atttttctaa atacattcaa atatgtatcc gctcatgaga   4140
caataaccct gataaatgct tcaataatat tgaaaaagga agagtatgag tattcaacat   4200
ttccgtgtcg cccttattcc ctttttgcg gcattttgcc ttcctgtttt tgctcaccca   4260
gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg gtgcacgagt gggttacatc   4320
gaactggatc tcaacagcgg taagatcctt gagagttttc gccccgaaga acgttttcca   4380
atgatgagca cttttaaagt tctgctatgt ggcgcggtat tatcccgtat tgacgccggg   4440
caagagcaac tcggtcgccg catacactat tctcagaatg acttggttga gtactcacca   4500
gtcacagaaa agcatcttac ggatggcatg acagtaagag aattatgcag tgctgccata   4560
accatgagtg ataacactgc ggccaactta cttctgacaa cgatcggagg accgaaggag   4620
ctaaccgctt ttttgcacaa catgggggat catgtaactc gccttgatcg ttgggaaccg   4680
gagctgaatg aagccatacc aaacgacgag cgtgacacca cgatgcctgt agcaatggca   4740
acaacgttgc gcaaactatt aactggcgaa ctacttactc tagcttcccg gcaacaatta   4800
atagactgga tggaggcgga taaagttgca ggaccacttc tgcgctcggc ccttccggct   4860
ggctggttta ttgctgataa atctggagcc ggtgagcgtg gttctcgcgg tatcattgca   4920
gcactggggc cagatggtaa gccctcccgt atcgtagtta tctacacgac ggggagtcag   4980
gcaactatgg atgaacgaaa tagacagatc gctgagatag gtgcctcact gattaagcat   5040
tggtaactgt cagaccaagt ttactcatat atactttaga ttgatttaaa acttcatttt   5100
taatttaaaa ggatctaggt gaagatcctt tttgataatc tcatgaccaa aatcccttaa   5160
cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga   5220
gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa aaaaaccacc gctaccagcg   5280
gtggtttgtt tgccggatca agagctacca actcttttc cgaaggtaac tggcttcagc   5340
agagcgcaga taccaaatac tgttcttcta gtgtagccgt agttaggcca ccacttcaag   5400
aactctgtag caccgcctac atacctcgct ctgctaatcc tgttaccagt ggctgctgcc   5460
agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg   5520
cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac   5580
accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga   5640
aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgcac gagggagctt   5700
ccagggggaa acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag   5760
cgtcgatttt tgtgatgctc gtcaggggggg cggagcctat ggaaaaacgc cagcaacgcg   5820
gccttttttac ggttcctggc cttttgctgg ccttttgctc acat                   5864
```

<210> 352

<211> 6428
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15796

<400> 352
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggcc aagtctggca tctaccagat taagaacacc 480
ctgaataaca aagtgtatgt tggatctgcc aaggactttg agaaaagatg gaagaggcac 540
ttcaaggatc tggaaaaggg ttgccacagc tccatcaagc tccaaaggtc tttcaacaaa 600
catggcaatg tctttgagtg ctccattttg gaagagatcc cttatgaaaa ggacctcatt 660
atagagaggg aaaacttctg gatcaaggag ctgaactcaa aaatcaatgg atacaacatt 720
gcagatgcca catttggtga cacctgttct actcaccccc ttaaggaaga gatcattaaa 780
aagaggtctg agactgtgaa ggctaagatg ttgaaactgg ggcctgatgg cagaaaggca 840
ctctactcta agccaggtag caaaaatgga aggtggaacc ctgagaccca caagttctgc 900
aaatgtgggg tgagaataca gacatcagcc tacacttgca gcaagtgtag gaacagatca 960
ggagagaata actccttttt caaccacaag cattctcaag gtccatccgg agtggatcta 1020
cgcacgctcg gctacagcca gcagcaacag gagaagatca aaccgaaggt tcgttcgaca 1080
gtggcgcagc accacgaggc actggtcggc cacgggttta cacacgcgca catcgttgcg 1140
ttaagccaac acccggcagc gttagggacc gtcgctgtca agtatcagga catgatcgca 1200
gcgttgccag aggcgacaca cgaagcgatc gttggcgtcg gcaaacagtg gtccggcgca 1260
cgcgctctgg aggccttgct cacggtggcg ggagagttga gaggtccacc gttacagttg 1320
gacacaggcc aacttctcaa gattgcaaaa cgtggcggcg tgaccgcagt ggaggcagtg 1380
catgcatggc gcaatgcact gacgggtgcc ccgctcaact tgaccccgga gcaggtggtg 1440
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 1500
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 1560
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 1620
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg 1680
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 1740
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 1800
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 1860
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 1920
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag 1980
gcgctggaga cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 2040
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 2100
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 2160
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 2220
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 2280
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 2340
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 2400
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 2460
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 2520
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagcaatatt 2580
ggtggcaagc aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac 2640
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg 2700
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag 2760
gtggtggcca tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg 2820
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 2880
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2940
gcccacggct tgaccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 3000
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 3060
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 3120
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc ctcagcaggt ggtggccatc 3180
gccagcaatg gcggcggcag gccggcgctg gagagcattg ttgcccagtt atctcgccct 3240
gatccgggat cctccggacc gaacagaggc gttaccaagc aactagttaa atccgaactg 3300
gaagagaaaa agtccgagtt gcgtcataaa ctgaagtacg ttccacatga atacattgaa 3360

ctgatcgaga ttgcccgtaa cagcacccag gacagaatcc tggaaatgaa agtgatggag 3420
tttttcatga aagtgtacgg ctacagaggc aaacatttgg gcggtagccg taaaccggat 3480
ggtgccatct acaccgttgg cagcccgatc gattacggcg tgatcgttga taccaaagcc 3540
tactccggtg gctacaacct gccgattggc caggctgacg aaatgcagag atacgtggaa 3600
gagaaccaga ctcgtaacaa gcatatcaac ccgaacgaat ggtggaaagt ctacccgagt 3660
tccgtgactg agttcaaatt tctgtttgtg agcggccatt tcaaaggcaa ctacaaagcc 3720
cagctgacca gactgaacca tatcaccaac tgcaacggtg ccgtcttgag cgtggaagaa 3780
ctgttgattg gtggcgaaat gatcaaggct ggcaccctga ctctggagga agttagacgt 3840
aaattcaaca acggtgaaat caactttgga tcctcggccg actgataagt cgaccaccac 3900
caccaccacc actgataaga gctcctcgag cctgcagcag ctgaagcttt ggacttctgg 3960
cgcgcccaat tcgccctata gtgagtcgta ttacgtcgcg ctcactggcc gtcgttttac 4020
aacgtcgtga ctgggaaaac cctggcgtta cccaacttaa tcgccttgca gcacatcccc 4080
ctttcgccag ctggcgtaat agcgaagagg cccgcaccga aacgcccttc ccaacagttg 4140
cgcagcctga atggcgaatg ggagcgccct gtagcggcgc attaagcgcg gcgggtgtgg 4200
tggttacgcg cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt 4260
tcttcccttc ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc 4320
tccctttagg gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg 4380
gtgatggttg gcctgtagtg ggccatagcc ctgatagacg gtttttcgcc ctttgacgtt 4440
ggagtccacg ttctttaata gtggactctt gttccaaact ggaacaacac tcaaccctat 4500
ctcggtctat tcttttgatt tataagggat tttgccgatt tcggcctatt ggttaaaaaa 4560
tgagctgatt taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc ttacaattta 4620
ggtggcactt ttcggggaaa tgtgcgcgga acccctattt gtttatttttt ctaaatacat 4680
tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa 4740
aggaagagta tgagtattca acatttccgt gtcgccctta ttccctttttt tgcggcattt 4800
tgccttcctg tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag 4860
ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaagat ccttgagagt 4920
tttcgccccg aagaacgttt tccaatgatg agcacttttа aagttctgct atgtggcgcg 4980
gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag 5040
aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta 5100
agagaattat gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg 5160
acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta 5220
actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac 5280
accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt 5340
actctagctt cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca 5400
cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag 5460
cgtggttctc gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta 5520
gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag 5580
ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt 5640
tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat ccttttttgat 5700
aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta 5760
gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa 5820
acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt 5880
tttccgaagg taactggctt cagcagagcg cagataccaa atactgttct tctagtgtag 5940
ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta 6000
atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca 6060
agacgatagt taccggataa ggcgcagcgg tcgggctgaa cgggggttc gtgcacacag 6120
cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa 6180
agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga 6240
acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc 6300
gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc 6360
ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt 6420
gctcacat 6428

<210> 353

<211> 6362
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15797

<400> 353
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240

ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggcc aagtctggca tctaccagat taagaacacc 480
ctgaataaca aagtgtatgt tggatctgcc aaggactttg agaaaagatg gaagaggcac 540
ttcaaggatc tggaaaaggg ttgccacagc tccatcaagc tccaaaggtc tttcaacaaa 600
catggcaatg tctttgagtg ctccattttg gaagagatcc cttatgaaaa ggacctcatt 660
atagagaggg aaaacttctg gatcaaggag ctgaactcaa aaatcaatgg atacaacatt 720
gcagatgcca catttggtga cacctgttct actcaccccc ttaaggaaga gatcattaaa 780
aagaggtctg agactgtgaa ggctaagatg ttgaaactgg ggcctgatgg cagaaaggca 840
ctctactcta agccaggtag caaaaatgga aggtggaacc ctgagaccca caagttctgc 900
aaatgtgggg tgagaataca gacatcagcc tacacttgca gcaagtgtag gaacagatca 960
ggagagaata actccttttt caaccacaag cattctcaag gtccatccgg agtggatcta 1020
cgcacgctcg gctacagcca gcagcaacag gagaagatca aaccgaaggt tcgttcgaca 1080
gtggcgcagc accacgaggc actggtcggc cacgggttta cacacgcgca catcgttgcg 1140
ttaagccaac acccggcagc gttagggacc gtcgctgtca agtatcagga catgatcgca 1200
gcgttgccag aggcgacaca cgaagcgatc gttggcgtcg gcaaacagtg gtccggcgca 1260
cgcgctctgg aggccttgct cacggtggcg ggagagttga gaggtccacc gttacagttg 1320
gacacaggcc aacttctcaa gattgcaaaa cgtggcggcg tgaccgcagt ggaggcagtg 1380
catgcatggc gcaatgcact gacgggtgcc ccgctcaact tgaccccgga gcaggtggtg 1440
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 1500
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 1560
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 1620
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg 1680
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 1740
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 1800
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 1860
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 1920
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag 1980
gcgctggaga cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 2040
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 2100
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 2160
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 2220
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 2280
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 2340
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 2400
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 2460
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 2520
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagcaatatt 2580
ggtggcaagc aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac 2640
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg 2700
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag 2760
gtggtggcca tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg 2820
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 2880
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2940
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 3000
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 3060
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 3120
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc ctcagcaggt ggtggccatc 3180
gccagcaatg gcggcggcag gccggcgctg gagagcattg ttgcccagtt atctcgccct 3240
gatccgggat ccaagtctgg catctaccag attaagaaca ccctgaataa caaagtgtat 3300
gttggatctg ccaaggactt tgagaaaaga tggaagaggc acttcaagga tctggaaaag 3360
ggttgccaca gctccatcaa gctccaaagg tctttcaaca aacatggcaa tgtctttgag 3420
tgctccattt tggaagagat cccttatgaa aaggacctca ttatagagag ggaaaacttc 3480
tggatcaagg agctgaactc aaaaatcaat ggatacaaca ttgcagatgc cacatttggt 3540
gacacctgtt ctactcaccc ccttaaggaa gagatcatta aaaagaggtc tgagactgtg 3600
aaggctaaga tgttgaaact ggggcctgat ggcagaaagg cactctactc taagccaggt 3660
agcaaaaatg gaaggtggaa ccctgagacc cacaagttct gcaaatgtgg ggtgagaata 3720
cagacatcag cctacacttg cagcaagtgt aggaacagat caggagagaa taactccttt 3780
ttcaaccaca agcattctca aggtccatcg gccgactgat aagtcgacca ccaccaccac 3840
caccactgat aagagctcct cgagcctgca gcagctgaag ctttggactt ctggcgcgcc 3900
caattcgccc tatagtgagt cgtattacgt cgcgctcact ggccgtcgtt ttacaacgtc 3960
gtgactggga aaaccctggc gttacccaac ttaatcgcct tgcagcacat ccccctttcg 4020
ccagctggcg taatagcgaa gaggcccgca ccgaaacgcc cttcccaaca gttgcgcagc 4080
ctgaatggcg aatgggagcg ccctgtagcg gcgcattaag cgcggcgggt gtggtggtta 4140
cgcgcagcgt gaccgctaca cttgccagcg ccctagcgcc cgctcctttc gctttcttcc 4200
cttcctttct cgccacgttc gccggctttc cccgtcaagc tctaaatcgg gggctccctt 4260
tagggttccg atttagtgct ttacggcacc tcgaccccaa aaaacttgat tagggtgatg 4320

gttggcctgt agtgggccat agccctgata gacggttttt cgccctttga cgttggagtc 4380
cacgttcttt aatagtggac tcttgttcca aactggaaca acactcaacc ctatctcggt 4440
ctattctttt gatttataag ggattttgcc gatttcggcc tattggttaa aaaatgagct 4500
gatttaacaa aaatttaacg cgaattttaa caaaatatta acgcttacaa tttaggtggc 4560
acttttcggg gaaatgtgcg cggaacccct atttgtttat ttttctaaat acattcaaat 4620
atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggaag 4680
agtatgagta ttcaacattt ccgtgtcgcc cttattccct tttttgcggc attttgcctt 4740
cctgtttttg ctcacccaga aacgctggtg aaagtaaaag atgctgaaga tcagttgggt 4800
gcacgagtgg gttacatcga actggatctc aacagcggta agatccttga gagttttcgc 4860
cccgaagaac gttttccaat gatgagcact tttaaagttc tgctatgtgg cgcggtatta 4920
tcccgtattg acgccgggca agagcaactc ggtcgccgca tacactattc tcagaatgac 4980
ttggttgagt actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa 5040
ttatgcagtg ctgccataac catgagtgat aacactgcgg ccaacttact tctgacaacg 5100
atcggaggac cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc 5160
cttgatcgtt gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg 5220
atgcctgtag caatggcaac aacgttgcgc aaactattaa ctggcgaact acttactcta 5280
gcttcccggc aacaattaat agactggatg gaggcggata aagttgcagg accacttctg 5340
cgctcggccc ttccggctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggt 5400
tctcgcggta tcattgcagc actggggcca gatggtaagc cctcccgtat cgtagttatc 5460
tacacgacgg ggagtcaggc aactatggat gaacgaaata gacagatcgc tgagataggt 5520
gcctcactga ttaagcattg gtaactgtca gaccaagttt actcatatat actttagatt 5580
gatttaaaac ttcattttta atttaaaagg atctaggtga agatcctttt tgataatctc 5640
atgaccaaaa tcccttaacg tgagttttcg ttccactgag cgtcagaccc cgtagaaaag 5700
atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa 5760
aaaccaccgc taccagcggt ggtttgtttg ccggatcaag agctaccaac tctttttccg 5820
aaggtaactg gcttcagcag agcgcagata ccaaatactg ttcttctagt gtagccgtag 5880
ttaggccacc acttcaagaa ctctgtagca ccgcctacat acctcgctct gctaatcctg 5940
ttaccagtgg ctgctgccag tggcgataag tcgtgtctta ccgggttgga ctcaagacga 6000
tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc 6060
ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagctatg agaaagcgcc 6120
acgcttcccg aagggagaaa ggcggacagg tatccggtaa gcggcagggt cggaacagga 6180
gagcgcacga gggagcttcc agggggaaac gcctggtatc tttatagtcc tgtcgggttt 6240
cgccacctct gacttgagcg tcgatttttg tgatgctcgt caggggggcg gagcctatgg 6300
aaaaacgcca gcaacgcggc ctttttacgg ttcctggcct tttgctggcc ttttgctcac 6360
at 6362

<210> 354

<211> 11739
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS13299

<400> 354
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
tttttgtttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260

atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg ccaagtctgg catctaccag 2460
attaagaaca ccctgaataa caaagtgtat gttggatctg ccaaggactt tgagaaaaga 2520
tggaagaggc acttcaagga tctggaaaag ggttgccaca gctccatcaa gctccaaagg 2580
tctttcaaca aacatggcaa tgtctttgag tgctccattt tggaagagat cccttatgaa 2640
aaggacctca ttatagagag ggaaaacttc tggatcaagg agctgaactc aaaaatcaat 2700
ggatacaaca ttgcagatgc cacatttggt gacacctgtt ctactcaccc ccttaaggaa 2760
gagatcatta aaaagaggtc tgagactgtg aaggctaaga tgttgaaact ggggcctgat 2820
ggcagaaagg cactctactc taagccaggt agcaaaaatg gaaggtggaa ccctgagacc 2880
cacaagttct gcaaatgtgg ggtgagaata cagacatcag cctacacttg cagcaagtgt 2940
aggaacagat caggagagaa taactccttt ttcaaccaca agcattctca aggtccatcc 3000
ggagtggatc tacgcacgct cggctacagc cagcagcaac aggagaagat caaaccgaag 3060
gttcgttcga cagtggcgca gcaccacgag gcactggtcg gccacgggtt tacacacgcg 3120
cacatcgttg cgttaagcca acacccggca gcgttaggga ccgtcgctgt caagtatcag 3180
gacatgatcg cagcgttgcc agaggcgaca cacgaagcga tcgttggcgt cggcaaacag 3240
tggtccggcg cacgcgctct ggaggccttg ctcacggtgg cgggagagtt gagaggtcca 3300
ccgttacagt tggacacagg ccaacttctc aagattgcaa aacgtggcgg cgtgaccgca 3360
gtggaggcag tgcatgcatg gcgcaatgca ctgacgggtg ccccgctcaa cttgaccccg 3420
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 3480
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 3540
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3600
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 3660
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 3720
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 3780
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 3840
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 3900
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagcaatatt 3960
ggtggcaagc aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac 4020
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg 4080
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag 4140
gtggtggcca tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg 4200
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 4260
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 4320
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 4380
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 4440
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 4500
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 4560
gccagcaata ttggtggcaa gcaggcgctg gagacggtgc aggcgctgtt gccggtgctg 4620
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 4680
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 4740
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 4800
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 4860
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 4920
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 4980
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 5040
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg 5100
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac ccctcagcag 5160
gtggtggcca tcgccagcaa tggcggcggc aggccggcgc tggagagcat tgttgcccag 5220
ttatctcgcc ctgatccggg atcctccgga ccgaacagag gcgttaccaa gcaactagtt 5280
aaatccgaac tggaagagaa aaagtccgag ttgcgtcata aactgaagta cgttccacat 5340

gaatacattg aactgatcga gattgcccgt aacagcaccc aggacagaat cctggaaatg 5400
aaagtgatgg agtttttcat gaaagtgtac ggctacagag gcaaacattt gggcggtagc 5460
cgtaaaccgg atggtgccat ctacaccgtt ggcagcccga tcgattacgg cgtgatcgtt 5520
gataccaaag cctactccgg tggctacaac ctgccgattg gccaggctga cgaaatgcag 5580
agatacgtgg aagagaacca gactcgtaac aagcatatca acccgaacga atggtggaaa 5640
gtctacccga gttccgtgac tgagttcaaa tttctgtttg tgagcggcca tttcaaaggc 5700
aactacaaag cccagctgac cagactgaac catatcacca actgcaacgg tgccgtcttg 5760
agcgtggaag aactgttgat tggtggcgaa atgatcaagg ctggcaccct gactctggag 5820
gaagttagac gtaaattcaa caacggtgaa atcaactttg gatcctcggc cgactgataa 5880
ctcgagcgat cctctagacg agctcctcga gcctgcagca gctgaagctt tggacttctt 5940
cgccagaggt ttggtcaagt ctccaatcaa ggttgtcggc ttgtctacct tgccagaaat 6000
ttacgaaaag atggaaaagg gtcaaatcgt tggtagatac gttgttgaca cttctaaata 6060
agcgaatttc ttatgattta tgatttttat tattaaataa gttataaaaa aaataagtgt 6120
atacaaattt taaagtgact cttaggtttt aaaacgaaaa ttcttattct tgagtaactc 6180
tttcctgtag gtcaggttgc tttctcaggt atagcatgag gtcgctctta ttgaccacac 6240
ctctaccggc atgcaagctt ggcgtaatca tggtcatagc tgtttcctgt gtgaaattgt 6300
tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa agcctggggt 6360
gcctaatgag tgagctaact cacattaatt gcgttgcgct cactgcccgc tttccagtcg 6420
ggaaacctgt cgtgccagca gatctattac attatgggtg gtatgttgga ataaaaatca 6480
actatcatct actaactagt atttacgtta ctagtatatt atcatatacg gtgttagaag 6540
atgacgcaaa tgatgagaaa tagtcatcta aattagtgga agctgaaacg caaggattga 6600
taatgtaata ggatcaatga atattaacat ataaaatgat gataataata tttatagaat 6660
tgtgtagaat tgcagattcc cttttatgga ttcctaaatc ctcgaggaga acttctagta 6720
tatctacata cctaatatta ttgccttatt aaaaatggaa tcccaacaat tacatcaaaa 6780
tccacattct cttcaaaatc aattgtcctg tacttccttg ttcatgtgtg ttcaaaaacg 6840
ttatatttat aggataatta tactctattt ctcaacaagt aattggttgt ttggccgagc 6900
ggtctaaggc gcctgattca agaaatatct tgaccgcagt taactgtggg aatactcagg 6960
tatcgtaaga tgcaagagtt cgaatctctt agcaaccatt atttttttcc tcaacataac 7020
gagaacacac aggggcgcta tcgcacagaa tcaaattcga tgactggaaa ttttttgtta 7080
atttcagagg tcgcctgacg catatacctt tttcaactga aaaattggga gaaaaaggaa 7140
aggtgagagc cgcggaaccg gcttttcata tagaatagag aagcgttcat gactaaatgc 7200
ttgcatcaca atacttgaag ttgacaatat tatttaagga cctattgttt tttccaatag 7260
gtggttagca atcgtcttac tttctaactt ttcttacctt ttacatttca gcaatatata 7320
tatatatatt tcaaggatat accattctaa tgtctgcccc taagaagatc gtcgttttgc 7380
caggtgacca cgttggtcaa gaaatcacag ccgaagccat taaggttctt aaagctattt 7440
ctgatgttcg ttccaatgtc aagttcgatt tcgaaaatca tttaattggt ggtgctgcta 7500
tcgatgctac aggtgtccca cttccagatg aggcgctgga agcctccaag aaggttgatg 7560
ccgttttgtt aggtgctgtg ggtggtccta aatggggtac cggtagtgtt agacctgaac 7620
aaggtttact aaaaatccgt aaagaacttc aattgtacgc caacttaaga ccatgtaact 7680
ttgcatccga ctctctttta gacttatctc caatcaagcc acaatttgct aaaggtactg 7740
acttcgttgt tgtcagagaa ttagtgggag gtatttactt tggtaagaga aaggaagacg 7800
atggtgatgg tgtcgcttgg gatagtgaac aatacaccgt tccagaagtg caaagaatca 7860
caagaatggc cgctttcatg gccctacaac atgagccacc attgcctatt tggtccttgg 7920
ataaagctaa tgttttggcc tcttcaagat tatggagaaa aactgtggag gaaaccatca 7980
agaacgaatt ccctacattg aaggttcaac atcaattgat tgattctgcc gccatgatcc 8040
tagttaagaa cccaacccac ctaaatggta ttataatcac cagcaacatg tttggtgata 8100
tcatctccga tgaagcctcc gttatcccag gttccttggg tttgttgcca tctgcgtcct 8160
tggcctcttt gccagacaag aacaccgcat ttggtttgta cgaaccatgc cacggttctg 8220
ctccagattt gccaaagaat aaggtcaacc ctatcgccac tatcttgtct gctgcaatga 8280
tgttgaaatt gtcattgaac ttgcctgaag aaggtaaggc cattgaagat gcagttaaaa 8340
aggttttgga tgcaggtatc agaactggtg atttaggtgg ttccaacagt accacggaag 8400
tcggtgatgc tgtcgccgaa gaagttaaga aaatccttgc ttaaaaagat tctctttttt 8460
tatgatattt gtacataaac tttataaatg aaattcataa tagaaacgac acgaaattac 8520
aaaatggaat atgttcatag ggtagacgaa actatatacg caatctacat acatttatca 8580
agaaggagaa aaaggaggat gtaaaggaat acaggtaagc aaattgatac taatggctca 8640
acgtgataag gaaaaagaat tgcactttaa cattaatatt gacaaggagg agggcaccac 8700
acaaaaagtt aggtgtaaca gaaaatcatg aaactatgat tcctaattta tatattggag 8760
gattttctct aaaaaaaaaa aaatacaaca aataaaaaac actcaatgac ctgaccattt 8820
gatggagttt aagtcaatac cttcttgaac catttcccat aatggtgaaa gttccctcaa 8880
gaattttact ctgtcagaaa cggccttaac gacgtagtcg acctcctctt cagtactaaa 8940
tctaccaata ccaaatctga tggaagaatg ggctaatgca tcatccttac ccagcgcatg 9000
taaaacataa gaaggttcta gggaagcaga tgtacaggct gaacccgagg ataatgcgat 9060
atcccttagt gccatcaata aagattctcc ttccacgtag gcgaaagaaa cgttaacaca 9120
ccctggataa cgatgatctg gagatccgtt caacgtggta tgttcagcgg ataatagacc 9180
tttgactaat ttatcggata gtcttttgat gtgagcttgg tcgttgtcaa attctttctt 9240
catcaatctc gcagcttcac caaatcccgc taccaatggg ggggccaaag taccagatct 9300
gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg ggcgctcttc 9360
cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc 9420

atcgatgaat tccacggact atagactata ctagtatact ccgtctactg tacgatacac 9480
ttccgctcag gtccttgtcc tttaacgagg ccttaccact cttttgttac tctattgatc 9540
cagctcagca aaggcagtgt gatctaagat tctatcttcg cgatgtagta aaactagcta 9600
gaccgagaaa gagactagaa atgcaaaagg cacttctaca atggctgcca tcattattat 9660
ccgatgtgac gctgcagctt ctcaatgata ttcgaatacg ctttgaggag atacagccta 9720
atatccgaca aactgtttta cagatttacg atcgtacttg ttacccatca ttgaattttg 9780
aacatccgaa cctgggagtt ttccctgaaa cagatagtat atttgaacct gtataataat 9840
atatagtcta gcgctttacg gaagacaatg tatgtatttc ggttcctgga gaaactattg 9900
catctattgc ataggtaatc ttgcacgtcg catccccggt tcattttctg cgtttccatc 9960
ttgcacttca atagcatatc tttgttaacg aagcatctgt gcttcatttt gtagaacaaa 10020
aatgcaacgc gagagcgcta atttttcaaa caaagaatct gagctgcatt tttacagaac 10080
agaaatgcaa cgcgaaagcg ctattttacc aacgaagaat ctgtgcttca tttttgtaaa 10140
acaaaaatgc aacgcgagag cgctaatttt tcaaacaaag aatctgagct gcatttttac 10200
agaacagaaa tgcaacgcga gagcgctatt ttaccaacaa agaatctata cttctttttt 10260
gttctacaaa aatgcatccc gagagcgcta tttttctaac aaagcatctt agattacttt 10320
ttttctcctt tgtgcgctct ataatgcagt ctcttgataa ctttttgcac tgtaggtccg 10380
ttaaggttag aagaaggcta ctttggtgtc tattttctct tccataaaaa aagcctgact 10440
ccacttcccg cgtttactga ttactagcga agctgcgggt gcatttttc aagataaagg 10500
catccccgat tatattctat accgatgtgg attgcgcata ctttgtgaac agaaagtgat 10560
agcgttgatg attcttcatt ggtcagaaaa ttatgaacgg tttcttctat tttgtctcta 10620
tatactacgt ataggaaatg tttacatttt cgtattgttt tcgattcact ctatgaatag 10680
ttcttactac aatttttttg tctaaagagt aatactagag ataaacataa aaaatgtaga 10740
ggtcgagttt agatgcaagt tcaaggagcg aaaggtggat gggtaggtta tatagggata 10800
tagcacagag atatatagca aagagatact tttgagcaat gtttgtggaa gcggtattcg 10860
caatatttta gtagctcgtt acagtccggt gcgtttttgg ttttttgaaa gtgcgtcttc 10920
agagcgcttt tggttttcaa aagcgctctg aagttcctat actttctaga gaataggaac 10980
ttcggaatag gaacttcaaa gcgtttccga aaacgagcgc ttccgaaaat gcaacgcgag 11040
ctgcgcacat acagctcact gttcacgtcg cacctatatc tgcgtgttgc ctgtatatat 11100
atatacatga gaagaacggc atagtgcgtg tttatgctta aatgcgtact tatatgcgtc 11160
tatttatgta ggatgaaagg tagtctagta cctcctgtga tattatccca ttccatgcgg 11220
ggtatcgtat gcttccttca gcactaccct ttagctgttc tatatgctgc cactcctcaa 11280
ttggattagt ctcatccttc aatgctatca tttcctttga tattggatca tatgcatagt 11340
accgagaaac tagtgcgaag tagtgatcag gtattgctgt tatctgatga gtatacgttg 11400
tcctggccac ggcagaagca cgcttatcgc tccaatttcc cacaacatta gtcaactccg 11460
ttaggccctt cattgaaaga aatgaggtca tcaaatgtct tccaatgtga gattttgggc 11520
catttttat agcaaagatt gaataaggcg cattttctt caaagcttta ttgtacgatc 11580
tgactaagtt atcttttaat aattggtatt cctgtttatt gcttgaagaa ttgccggtcc 11640
tatttactcg ttttaggact ggttcagaat tcatcgatgc tcactcaaag gtcggtaata 11700
cggttatcca cagaatcagg ggataacgca ggaaagaac 11739

<210> 355

<211> 11673
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS13301

<400> 355
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020

cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa atagggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg ccaagtctgg catctaccag 2460
attaagaaca ccctgaataa caaagtgtat gttggatctg ccaaggactt tgagaaaaga 2520
tggaagaggc acttcaagga tctggaaaag ggttgccaca gctccatcaa gctccaaagg 2580
tctttcaaca aacatggcaa tgtctttgag tgctccattt tggaagagat cccttatgaa 2640
aaggacctca ttatagagag ggaaaacttc tggatcaagg agctgaactc aaaaatcaat 2700
ggatacaaca ttgcagatgc cacatttggt gacacctgtt ctactcaccc ccttaaggaa 2760
gagatcatta aaaagaggtc tgagactgtg aaggctaaga tgttgaaact ggggcctgat 2820
ggcagaaagg cactctactc taagccaggt agcaaaaatg gaaggtggaa ccctgagacc 2880
cacaagttct gcaaatgtgg ggtgagaata cagacatcag cctacacttg cagcaagtgt 2940
aggaacagat caggagagaa taactccttt ttcaaccaca agcattctca aggtccatcc 3000
ggagtggatc tacgcacgct cggctacagc cagcagcaac aggagaagat caaaccgaag 3060
gttcgttcga cagtggcgca gcaccacgag gcactggtcg gccacgggtt tacacacgcg 3120
cacatcgttg cgttaagcca acacccggca gcgttaggga ccgtcgctgt caagtatcag 3180
gacatgatcg cagcgttgcc agaggcgaca cacgaagcga tcgttggcgt cggcaaacag 3240
tggtccggcg cacgcgctct ggaggccttg ctcacggtgg cgggagagtt gagaggtcca 3300
ccgttacagt tggacacagg ccaacttctc aagattgcaa aacgtggcgg cgtgaccgca 3360
gtggaggcag tgcatgcatg gcgcaatgca ctgacgggtg ccccgctcaa cttgaccccg 3420
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 3480
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 3540
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3600
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 3660
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 3720
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 3780
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 3840
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 3900
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagcaatatt 3960
ggtggcaagc aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac 4020
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg 4080
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag 4140
gtggtggcca tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg 4200
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 4260
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 4320
gcccacggct tgaccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 4380
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 4440
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 4500
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 4560
gccagcaata ttggtggcaa gcaggcgctg gagacggtgc aggcgctgtt gccggtgctg 4620
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 4680
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 4740
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 4800
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 4860
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 4920
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 4980
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 5040
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg 5100

gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac ccctcagcag 5160
gtggtggcca tcgccagcaa tggcggcggc aggccggcgc tggagagcat tgttgcccag 5220
ttatctcgcc ctgatccggg atccaagtct ggcatctacc agattaagaa caccctgaat 5280
aacaaagtgt atgttggatc tgccaaggac tttgagaaaa gatggaagag gcacttcaag 5340
gatctggaaa agggttgcca cagctccatc aagctccaaa ggtctttcaa caaacatggc 5400
aatgtctttg agtgctccat tttggaagag atcccttatg aaaaggacct cattatagag 5460
agggaaaact tctggatcaa ggagctgaac tcaaaaatca atggatacaa cattgcagat 5520
gccacatttg gtgacacctg ttctactcac ccccttaagg aagagatcat taaaaagagg 5580
tctgagactg tgaaggctaa gatgttgaaa ctggggcctg atggcagaaa ggcactctac 5640
tctaagccag gtagcaaaaa tggaaggtgg aaccctgaga cccacaagtt ctgcaaatgt 5700
ggggtgagaa tacagacatc agcctacact tgcagcaagt gtaggaacag atcaggagag 5760
aataactcct ttttcaacca caagcattct caaggtccat cggccgactg ataactcgag 5820
cgatcctcta gacgagctcc tcgagcctgc agcagctgaa gctttggact tcttcgccag 5880
aggtttggtc aagtctccaa tcaaggttgt cggcttgtct accttgccag aaatttacga 5940
aaagatggaa aagggtcaaa tcgttggtag atacgttgtt gacacttcta aataagcgaa 6000
tttcttatga tttatgattt ttattattaa ataagttata aaaaaaataa gtgtatacaa 6060
attttaaagt gactcttagg ttttaaaacg aaaattctta ttcttgagta actctttcct 6120
gtaggtcagg ttgctttctc aggtatagca tgaggtcgct cttattgacc acacctctac 6180
cggcatgcaa gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg 6240
ctcacaattc cacacaacat acgagccgga agcataaagt gtaaagcctg gggtgcctaa 6300
tgagtgagct aactcacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac 6360
ctgtcgtgcc agcagatcta ttacattatg ggtggtatgt tggaataaaa atcaactatc 6420
atctactaac tagtatttac gttactagta tattatcata tacggtgtta gaagatgacg 6480
caaatgatga gaaatagtca tctaaattag tggaagctga aacgcaagga ttgataatgt 6540
aataggatca atgaatatta acatataaaa tgatgataat aatatttata gaattgtgta 6600
gaattgcaga ttccctttta tggattccta aatcctcgag gagaacttct agtatatcta 6660
catacctaat attattgcct tattaaaaat ggaatcccaa caattacatc aaaatccaca 6720
ttctcttcaa aatcaattgt cctgtacttc cttgttcatg tgtgttcaaa aacgttatat 6780
ttataggata attatactct atttctcaac aagtaattgg ttgtttggcc gagcggtcta 6840
aggcgcctga ttcaagaaat atcttgaccg cagttaactg tgggaatact caggtatcgt 6900
aagatgcaag agttcgaatc tcttagcaac cattattttt ttcctcaaca taacgagaac 6960
acacaggggc gctatcgcac agaatcaaat tcgatgactg gaaatttttt gttaatttca 7020
gaggtcgcct gacgcatata ccttttttcaa ctgaaaaatt gggagaaaaa ggaaaggtga 7080
gagccgcgga accggctttt catatagaat agagaagcgt tcatgactaa atgcttgcat 7140
cacaatactt gaagttgaca atattattta aggacctatt gtttttttca ataggtggtt 7200
agcaatcgtc ttactttcta acttttctta ccttttacat ttcagcaata tatatatata 7260
tatttcaagg atataccatt ctaatgtctg cccctaagaa gatcgtcgtt ttgccaggtg 7320
accacgttgg tcaagaaatc acagccgaag ccattaaggt tcttaaagct atttctgatg 7380
ttcgttccaa tgtcaagttc gatttcgaaa atcatttaat tggtggtgct gctatcgatg 7440
ctacaggtgt cccacttcca gatgaggcgc tggaagcctc caagaaggtt gatgccgttt 7500
tgttaggtgc tgtgggtggt cctaaatggg gtaccggtag tgttagacct gaacaaggtt 7560
tactaaaaat ccgtaaagaa cttcaattgt acgccaactt aagaccatgt aactttgcat 7620
ccgactctct tttagactta tctccaatca agccacaatt tgctaaaggt actgacttcg 7680
ttgttgtcag agaattagtg ggaggtattt actttggtaa gagaaaggaa gacgatggtg 7740
atggtgtcgc ttgggatagt gaacaataca ccgttccaga agtgcaaaga atcacaagaa 7800
tggccgcttt catggcccta caacatgagc caccattgcc tatttggtcc ttggataaag 7860
ctaatgtttt ggcctcttca agattatgga gaaaaactgt ggaggaaacc atcaagaacg 7920
aattccctac attgaaggtt caacatcaat tgattgattc tgccgccatg atcctagtta 7980
agaacccaac ccacctaaat ggtattataa tcaccagcaa catgtttggt gatatcatct 8040
ccgatgaagc ctccgttatc ccaggttcct tgggtttgtt gccatctgcg tccttggcct 8100
ctttgccaga caagaacacc gcatttggtt tgtacgaacc atgccacggt tctgctccag 8160
atttgccaaa gaataaggtc aaccctatcg ccactatctt gtctgctgca atgatgttga 8220
aattgtcatt gaacttgcct gaagaaggta aggccattga agatgcagtt aaaaaggttt 8280
tggatgcagg tatcagaact ggtgatttag gtggttccaa cagtaccacg gaagtcggtg 8340
atgctgtcgc cgaagaagtt aagaaaatcc ttgcttaaaa agattctctt tttttatgat 8400
atttgtacat aaactttata aatgaaattc ataatagaaa cgacacgaaa ttacaaaatg 8460
gaatatgttc atagggtaga cgaaactata tacgcaatct acatacattt atcaagaagg 8520
agaaaaagga ggatgtaaag gaatacaggt aagcaaattg atactaatgg ctcaacgtga 8580
taaggaaaaa gaattgcact ttaacattaa tattgacaag gaggagggca ccacacaaaa 8640
agttaggtgt aacagaaaat catgaaacta tgattcctaa tttatatatt ggaggatttt 8700
ctctaaaaaa aaaaaaatac aacaaataaa aaacactcaa tgacctgacc atttgatgga 8760
gtttaagtca ataccttctt gaaccatttc ccataatggt gaaagttccc tcaagaattt 8820
tactctgtca gaaacggcct taacgacgta gtcgacctcc tcttcagtac taaatctacc 8880
aataccaaat ctgatggaag aatgggctaa tgcatcatcc ttacccagcg catgtaaaac 8940
ataagaaggt tctagggaag cagatgtaca ggctgaaccc gaggataatg cgatatccct 9000
tagtgccatc aataaagatt ctccttccac gtaggcgaaa gaaacgttaa cacaccctgg 9060
ataacgatga tctggagatc cgttcaacgt ggtatgttca gcggataata gacctttgac 9120
taatttatcg gatagtcttt tgatgtgagc ttggtcgttg tcaaattctt tcttcatcaa 9180

tctcgcagct tcaccaaatc ccgctaccaa tgggggggcc aaagtaccag atctgctgca 9240
ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt attgggcgct cttccgcttc 9300
ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg cgagcggtat cagcatcgat 9360
gaattccacg gactatagac tatactagta tactccgtct actgtacgat acacttccgc 9420
tcaggtcctt gtcctttaac gaggccttac cactcttttg ttactctatt gatccagctc 9480
agcaaaggca gtgtgatcta agattctatc ttcgcgatgt agtaaaacta gctagaccga 9540
gaaagagact agaaatgcaa aaggcacttc tacaatggct gccatcatta ttatccgatg 9600
tgacgctgca gcttctcaat gatattcgaa tacgctttga ggagatacag cctaatatcc 9660
gacaaactgt tttacagatt tacgatcgta cttgttaccc atcattgaat tttgaacatc 9720
cgaacctggg agttttccct gaaacagata gtatatttga acctgtataa taatatatag 9780
tctagcgctt tacggaagac aatgtatgta tttcggttcc tggagaaact attgcatcta 9840
ttgcataggt aatcttgcac gtcgcatccc cggttcattt tctgcgtttc catcttgcac 9900
ttcaatagca tatctttgtt aacgaagcat ctgtgcttca ttttgtagaa caaaaatgca 9960
acgcgagagc gctaattttt caaacaaaga atctgagctg cattttaca gaacagaaat 10020
gcaacgcgaa agcgctattt taccaacgaa gaatctgtgc ttcatttttg taaaacaaaa 10080
atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt ttacagaaca 10140
gaaatgcaac gcgagagcgc tattttacca acaaagaatc tatacttctt ttttgttcta 10200
caaaaatgca tcccgagagc gctatttttc taacaaagca tcttagatta ctttttttct 10260
cctttgtgcg ctctataatg cagtctcttg ataacttttt gcactgtagg tccgttaagg 10320
ttagaagaag gctactttgg tgtctatttt ctcttccata aaaaaagcct gactccactt 10380
cccgcgttta ctgattacta gcgaagctgc gggtgcattt tttcaagata aaggcatccc 10440
cgattatatt ctataccgat gtggattgcg catactttgt gaacagaaag tgatagcgtt 10500
gatgattctt cattggtcag aaaattatga acggtttctt ctattttgtc tctatatact 10560
acgtatagga aatgtttaca ttttcgtatt gttttcgatt cactctatga atagttctta 10620
ctacaatttt tttgtctaaa gagtaatact agagataaac ataaaaaatg tagaggtcga 10680
gtttagatgc aagttcaagg agcgaaaggt ggatgggtag gttatatagg gatatagcac 10740
agagatatat agcaaagaga tacttttgag caatgtttgt ggaagcggta ttcgcaatat 10800
tttagtagct cgttacagtc cggtgcgttt ttggtttttt gaaagtgcgt cttcagagcg 10860
cttttggttt tcaaaagcgc tctgaagttc ctatactttc tagagaatag gaacttcgga 10920
ataggaactt caaagcgttt ccgaaaacga gcgcttccga aaatgcaacg cgagctgcgc 10980
acatacagct cactgttcac gtcgcaccta tatctgcgtg ttgcctgtat atatatatac 11040
atgagaagaa cggcatagtg cgtgtttatg cttaaatgcg tacttatatg cgtctattta 11100
tgtaggatga aaggtagtct agtacctcct gtgatattat cccattccat gcggggtatc 11160
gtatgcttcc ttcagcacta ccctttagct gttctatatg ctgccactcc tcaattggat 11220
tagtctcatc cttcaatgct atcatttcct ttgatattgg atcatatgca tagtaccgag 11280
aaactagtgc gaagtagtga tcaggtattg ctgttatctg atgagtatac gttgtcctgg 11340
ccacggcaga agcacgctta tcgctccaat ttcccacaac attagtcaac tccgttaggc 11400
ccttcattga aagaaatgag gtcatcaaat gtcttccaat gtgagatttt gggccatttt 11460
ttatagcaaa gattgaataa ggcgcatttt tcttcaaagc tttattgtac gatctgacta 11520
agttatcttt taataattgg tattcctgtt tattgcttga agaattgccg gtcctattta 11580
ctcgttttag gactggttca gaattcatcg atgctcactc aaaggtcggt aatacggtta 11640
tccacagaat caggggataa cgcaggaaag aac 11673

<210> 356

<211> 4967
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15798

<400> 356
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatgggt tccgaggcac cccgggccga gacctttgtc 480
ttcctggacc tggaagccac tgggctcccc agtgtggagc ccgagattgc cgagctgtcc 540
ctctttgctg tccaccgctc ctccctggag aacccggagc acgacgagtc tggtgcccta 600
gtattgcccc gggtcctgga caagctcacg ctgtgcatgt gcccggagcg ccccttcact 660
gccaaggcca gcgagatcac cggcctgagc agtgagggcc tggcgcgatg ccggaaggct 720
ggctttgatg gcgccgtggt gcggacgctg caggccttcc tgagccgcca ggcagggccc 780
atctgccttg tggcccacaa tggctttgat tatgatttcc ccctgctgtg tgccgagctg 840

cggcgcctgg gtgcccgcct gccccgggac actgtctgcc tggacacgct gccggccctg 900
cggggcctgg accgcgccca cagccacggc acccgggccc ggggccgcca gggttacagc 960
ctcggcagcc tcttccaccg ctacttccgg gcagagccaa gcgcagccca ctcagccgag 1020
ggcgacgtgc acaccctgct cctgatcttc ctgcaccgcg ccgcagagct gctcgcctgg 1080
gccgatgagc aggcccgtgg gtgggcccac atcgagccca tgtacttgcc gcctgatgac 1140
cccagcctgg aggcgactcc tccacagacc ggtctggatg ttccttactc cgaggcaccc 1200
cgggccgaga cctttgtctt cctggacctg gaagccactg ggctccccag tgtggagccc 1260
gagattgccg agctgtccct ctttgctgtc caccgctcct ccctggagaa cccggagcac 1320
gacgagtctg gtgccctagt attgccccgg gtcctggaca agctcacgct gtgcatgtgc 1380
ccggagcgcc ccttcactgc caaggccagc gagatcaccg gcctgagcag tgagggcctg 1440
gcgcgatgcc ggaaggctgg ctttgatggc gccgtggtgc ggacgctgca ggccttcctg 1500
agccgccagg cagggcccat ctgccttgtg gcccacaatg gctttgatta tgatttcccc 1560
ctgctgtgtg ccgagctgcg gcgcctgggt gcccgcctgc cccgggacac tgtctgcctg 1620
gacacgctgc cggccctgcg gggcctggac cgcgcccaca gccacggcac ccgggcccgg 1680
ggccgccagg gttacagcct cggcagcctc ttccaccgct acttccgggc agagccaagc 1740
gcagcccact cagccgaggg cgacgtgcac accctgctcc tgatcttcct gcaccgcgcc 1800
gcagagctgc tcgcctgggc cgatgagcag gcccgtgggt gggcccacat cgagcccatg 1860
tacttgccgc ctgatgaccc cagcctggag gcgggaggtg gaggttctgg aggtggagga 1920
tcctccggag tggatctacg cacgctcggc tacagccagc agcaacagga gaagatcaaa 1980
ccgaaggttc gttcgacagt ggcgcagcac cacgaggcac tggtcggcca cgggtttaca 2040
cacgcgcaca tcgttgcgtt aagccaacac ccggcagcgt tagggaccgt cgctgtcaag 2100
tatcaggaca tgatcgcagc gttgccagag gcgacacacg aagcgatcgt tggcgtcggc 2160
aaacagtggt ccggcgcacg cgctctggag gccttgctca cggtggcggg agagttgaga 2220
ggtccaccgt tacagttgga cacaggccaa cttctcaaga ttgcaaaacg tggcggcgtg 2280
accgcagtgg aggcagtgca tgcatggcgc aatgcactga cgggtgcccc gctcaacttg 2340
accggagacg cccgggggat caggtcacgt gcgtctcgga gcattgttgc ccagttatct 2400
cgccctgatc cgtcggccga ctgataagtc gaccaccacc accaccacca ctgataagag 2460
ctcctcgagc ctgcagcagc tgaagctttg gacttctggc gcgcccaatt cgccctatag 2520
tgagtcgtat tacgtcgcgc tcactggccg tcgttttaca acgtcgtgac tgggaaaacc 2580
ctggcgttac ccaacttaat cgccttgcag cacatccccc tttcgccagc tggcgtaata 2640
gcgaagaggc ccgcaccgaa acgcccttcc caacagttgc gcagcctgaa tggcgaatgg 2700
gagcgccctg tagcggcgca ttaagcgcgg cgggtgtggt ggttacgcgc agcgtgaccg 2760
ctacacttgc cagcgcccta gcgcccgctc ctttcgcttt cttcccttcc tttctcgcca 2820
cgttcgccgg ctttccccgt caagctctaa atcgggggct ccctttaggg ttccgattta 2880
gtgctttacg gcacctcgac cccaaaaaac ttgattaggg tgatggttgg cctgtagtgg 2940
gccatagccc tgatagacgg tttttcgccc tttgacgttg gagtccacgt tctttaatag 3000
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt 3060
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt 3120
taacgcgaat tttaacaaaa tattaacgct tacaatttag gtggcacttt tcggggaaat 3180
gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg 3240
agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa 3300
catttccgtg tcgcccttat tccctttttt gcggcatttt gccttcctgt ttttgctcac 3360
ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac 3420
atcgaactgg atctcaacag cggtaagatc cttgagagtt ttcgccccga agaacgtttt 3480
ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccg tattgacgcc 3540
gggcaagagc aactcggtcg ccgcatacac tattctcaga atgacttggt tgagtactca 3600
ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgctgcc 3660
ataaccatga gtgataacac tgcggccaac ttacttctga caacgatcgg aggaccgaag 3720
gagctaaccg ctttttgca caacatgggg gatcatgtaa ctcgccttga tcgttgggaa 3780
ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg 3840
gcaacaacgt tgcgcaaact attaactggc gaactactta ctctagcttc ccggcaacaa 3900
ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg 3960
gctggctggt ttattgctga taaatctgga gccggtgagc gtggttctcg cggtatcatt 4020
gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt 4080
caggcaacta tggatgaacg aaatagacag atcgctgaga taggtgcctc actgattaag 4140
cattggtaac tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat 4200
ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct 4260
taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct 4320
tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca 4380
gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt aactggcttc 4440
agcagagcgc agataccaaa tactgttctt ctagtgtagc cgtagttagg ccaccacttc 4500
aagaactctg tagcaccgcc tacatacctc gctctgctaa tcctgttacc agtggctgct 4560
gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag 4620
gcgcagcggt cgggctgaac gggggttcg tgcacacagc ccagcttgga gcgaacgacc 4680
tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg 4740
agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa caggagagcg cacgagggag 4800
cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt 4860
gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac 4920

gcggccttttt tacggttcct ggccttttgc tggccttttg ctcacat 4967

<210> 357

<211> 7346
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15799

<400> 357
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatgggt tccgaggcac cccgggccga gacctttgtc 480
ttcctggacc tggaagccac tgggctcccc agtgtggagc ccgagattgc cgagctgtcc 540
ctctttgctg tccaccgctc ctccctggag aacccggagc acgacgagtc tggtgcccta 600
gtattgcccc gggtcctgga caagctcacg ctgtgcatgt gcccggagcg ccccttcact 660
gccaaggcca gcgagatcac cggcctgagc agtgagggcc tggcgcgatg ccggaaggct 720
ggctttgatg gcgccgtggt gcggacgctg caggccttcc tgagccgcca ggcagggccc 780
atctgccttg tggcccacaa tggctttgat tatgatttcc ccctgctgtg tgccgagctg 840
cggcgcctgg gtgcccgcct gccccgggac actgtctgcc tggacacgct gccggccctg 900
cggggcctgg accgcgccca cagccacggc acccgggccc ggggccgcca gggttacagc 960
ctcggcagcc tcttccaccg ctacttccgg gcagagccaa gcgcagccca ctcagccgag 1020
ggcgacgtgc acaccctgct cctgatcttc ctgcaccgcg ccgcagagct gctcgcctgg 1080
gccgatgagc aggcccgtgg gtgggcccac atcgagccca tgtacttgcc gcctgatgac 1140
cccagcctgg aggcgactcc tccacagacc ggtctggatg ttccttactc cgaggcaccc 1200
cgggccgaga cctttgtctt cctggacctg gaagccactg ggctccccag tgtggagccc 1260
gagattgccg agctgtccct ctttgctgtc caccgctcct ccctggagaa cccggagcac 1320
gacgagtctg gtgccctagt attgccccgg gtcctggaca agctcacgct gtgcatgtgc 1380
ccggagcgcc ccttcactgc caaggccagc gagatcaccg gcctgagcag tgagggcctg 1440
gcgcgatgcc ggaaggctgg ctttgatggc gccgtggtgc ggacgctgca ggccttcctg 1500
agccgccagg cagggcccat ctgccttgtg gcccacaatg gctttgatta tgatttcccc 1560
ctgctgtgtg ccgagctgcg gcgcctgggt gcccgcctgc cccgggacac tgtctgcctg 1620
gacacgctgc cggccctgcg gggcctggac cgcgcccaca gccacggcac ccgggcccgg 1680
ggccgccagg gttacagcct cggcagcctc ttccaccgct acttccgggc agagccaagc 1740
gcagcccact cagccgaggg cgacgtgcac accctgctcc tgatcttcct gcaccgcgcc 1800
gcagagctgc tcgcctgggc cgatgagcag gcccgtgggt gggcccacat cgagcccatg 1860
tacttgccgc ctgatgaccc cagcctggag gcgggaggtg gaggttctgg aggtggagga 1920
tcctccggag tggatctacg cacgctcggc tacagccagc agcaacagga gaagatcaaa 1980
ccgaaggttc gttcgacagt ggcgcagcac cacgaggcac tggtcggcca cgggtttaca 2040
cacgcgcaca tcgttgcgtt aagccaacac ccggcagcgt tagggaccgt cgctgtcaag 2100
tatcaggaca tgatcgcagc gttgccagag gcgacacacg aagcgatcgt tggcgtcggc 2160
aaacagtggt ccggcgcacg cgctctggag gccttgctca cggtggcggg agagttgaga 2220
ggtccaccgt tacagttgga cacaggccaa cttctcaaga ttgcaaaacg tggcggcgtg 2280
accgcagtgg aggcagtgca tgcatggcgc aatgcactga cgggtgcccc gctcaacttg 2340
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 2400
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccca gcaggtggtg 2460
gccatcgcca gcaatggcgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg 2520
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagcaatatt 2580
ggtggcaagc aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac 2640
ggcttgaccc cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg 2700
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag 2760
gtggtggcca tcgccagcaa tattggtggc aagcaggcgc tggagacggt gcaggcgctg 2820
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 2880
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 2940
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag 3000
gcgctggaga cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 3060
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 3120
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 3180
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3240
tgccaggccc acggcttgac cccccagcag gtggtggcca tcgccagcaa tggcggtggc 3300

```
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg   3360
accccggagc aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg   3420
gtgcaggcgc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg   3480
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg   3540
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat   3600
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac   3660
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg   3720
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag   3780
gtggtggcca tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg   3840
ttgccggtgc tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc   3900
aatggcggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag   3960
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag   4020
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgacccct   4080
cagcaggtgg tggccatcgc cagcaatggc ggcggcaggc cggcgctgga gagcattgtt   4140
gcccagttat ctcgccctga tccgggatcc tccggaccga acagaggcgt taccaagcaa   4200
ctagttaaat ccgaactgga agagaaaaag tccgagttgc gtcataaact gaagtacgtt   4260
ccacatgaat acattgaact gatcgagatt gcccgtaaca gcacccagga cagaatcctg   4320
gaaatgaaag tgatggagtt tttcatgaaa gtgtacggct acagaggcaa acatttgggc   4380
ggtagccgta aaccggatgg tgccatctac accgttggca gcccgatcga ttacggcgtg   4440
atcgttgata ccaaagccta ctccggtggc tacaacctgc cgattggcca ggctgacgaa   4500
atgcagagat acgtggaaga gaaccagact cgtaacaagc atatcaaccc gaacgaatgg   4560
tggaaagtct acccgagttc cgtgactgag ttcaaatttc tgtttgtgag cggccatttc   4620
aaaggcaact acaaagccca gctgaccaga ctgaaccata tcaccaactg caacggtgcc   4680
gtcttgagcg tggaagaact gttgattggt ggcgaaatga tcaaggctgg caccctgact   4740
ctggaggaag ttagacgtaa attcaacaac ggtgaaatca actttggatc ctcggccgac   4800
tgataagtcg accaccacca ccaccaccac tgataagagc tcctcgagcc tgcagcagct   4860
gaagctttgg acttctggcg cgcccaattc gccctatagt gagtcgtatt acgtcgcgct   4920
cactggccgt cgttttacaa cgtcgtgact gggaaaaccc tggcgttacc caacttaatc   4980
gccttgcagc acatcccccct ttcgccagct ggcgtaatag cgaagaggcc cgcaccgaaa   5040
cgcccttccc aacagttgcg cagcctgaat ggcgaatggg agcgccctgt agcggcgcat   5100
taagcgcggc gggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc agcgccctag   5160
cgcccgctcc tttcgctttc ttcccttcct ttctcgccac gttcgccggc tttccccgtc   5220
aagctctaaa tcgggggctc cctttagggt tccgatttag tgctttacgg cacctcgacc   5280
ccaaaaaact tgattagggt gatggttggc ctgtagtggg ccatagccct gatagacggt   5340
ttttcgccct ttgacgttgg agtccacgtt ctttaatagt ggactcttgt tccaaactgg   5400
aacaacactc aaccctatct cggtctattc ttttgattta taagggattt tgccgatttc   5460
ggcctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt ttaacaaaat   5520
attaacgctt acaatttagg tggcactttt cggggaaatg tgcgcggaac ccctatttgt   5580
ttatttttct aaatacattc aaatatgtat ccgctcatga gacaataacc ctgataaatg   5640
cttcaataat attgaaaaag gaagagtatg agtattcaac atttccgtgt cgcccttatt   5700
ccctttttg cggcattttg ccttcctgtt tttgctcacc cagaaacgct ggtgaaagta   5760
aaagatgctg aagatcagtt gggtgcacga gtgggttaca tcgaactgga tctcaacagc   5820
ggtaagatcc ttgagagttt tcgccccgaa gaacgttttc caatgatgag cacttttaaa   5880
gttctgctat gtggcgcggt attatcccgt attgacgccg ggcaagagca actcggtcgc   5940
cgcatacact attctcagaa tgacttggtt gagtactcac cagtcacaga aaagcatctt   6000
acggatggca tgacagtaag agaattatgc agtgctgcca taaccatgag tgataacact   6060
gcggccaact tacttctgac aacgatcgga ggaccgaagg agctaaccgc ttttttgcac   6120
aacatggggg atcatgtaac tcgccttgat cgttgggaac cggagctgaa tgaagccata   6180
ccaaacgacg agcgtgacac cacgatgcct gtagcaatgg caacaacgtt gcgcaaacta   6240
ttaactggcg aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg   6300
gataaagttg caggaccact tctgcgctcg gcccttccgg ctggctggtt tattgctgat   6360
aaatctggag ccggtgagcg tggttctcgc ggtatcattg cagcactggg gccagatggt   6420
aagccctccc gtatcgtagt tatctacacg acggggagtc aggcaactat ggatgaacga   6480
aatagacaga tcgctgagat aggtgcctca ctgattaagc attggtaact gtcagaccaa   6540
gtttactcat atatacttta gattgattta aaacttcatt tttaatttaa aaggatctag   6600
gtgaagatcc tttttgataa tctcatgacc aaaatccctt aacgtgagtt ttcgttccac   6660
tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt ttttctgcgc   6720
gtaatctgct gcttgcaaac aaaaaaacca ccgctaccag cggtggtttg tttgccggat   6780
caagagctac caactctttt tccgaaggta actggcttca gcagagcgca gataccaaat   6840
actgttcttc tagtgtagcc gtagttaggc caccacttca agaactctgt agcaccgcct   6900
acatacctcg ctctgctaat cctgttacca gtggctgctg ccagtggcga taagtcgtgt   6960
cttaccgggt tggactcaag acgatagtta ccggataagg cgcagcggtc gggctgaacg   7020
gggggttcgt gcacacagcc cagcttggag cgaacgacct acaccgaact gagataccta   7080
cagcgtgagc tatgagaaag cgccacgctt cccgaaggga gaaaggcgga caggtatccg   7140
gtaagcggca gggtcggaac aggagagcgc acgagggagc ttccaggggg aaacgcctgg   7200
tatctttata gtcctgtcgg gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc   7260
tcgtcagggg ggcggagcct atggaaaaac gccagcaacg cggccttttt acggttcctg   7320
gccttttgct ggccttttgc tcacat                                       7346
```

<210> 358

<211> 10967
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS14972

<400> 358
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtttt 600
tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc 660
ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg 720
agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca 780
atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca 840
cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag 900
ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac 960
ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc 1020
agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct 1080
agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc 1140
gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg 1200
cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc 1260
gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat 1320
tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag 1380
tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat 1440
aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg 1500
cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca 1560
cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga 1620
aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc 1680
ttccttttc aatattattg aagcatttat cagggttatt gtctcatgag cggatacata 1740
tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 1800
ccacctgacg tcgacggatc gggagatctc ccgatcccct atggtgcact ctcagtacaa 1860
tctgctctga tgccgcatag ttaagccagt atctgctccc tgcttgtgtg ttggaggtcg 1920
ctgagtagtg cgcgagcaaa atttaagcta caacaaggca aggcttgacc gacaattgca 1980
tgaagaatct gcttagggtt aggcgttttg cgctgcttcg cgatgtacgg gccagatata 2040
cgcgttgaca ttgattattg actagttatt aatagtaatc aattacgggg tcattagttc 2100
atagcccata tatggagttc cgcgttacat aacttacggt aaatggcccg cctggctgac 2160
cgcccaacga cccccgccca ttgacgtcaa taatgacgta tgttcccata gtaacgccaa 2220
tagggacttt ccattgacgt caatgggtgg agtatttacg gtaaactgcc cacttggcag 2280
tacatcaagt gtatcatatg ccaagtacgc cccctattga cgtcaatgac ggtaaatggc 2340
ccgcctggca ttatgcccag tacatgacct tatgggactt tcctacttgg cagtacatct 2400
acgtattagt catcgctatt accatggtga tgcggttttg gcagtacatc aatgggcgtg 2460
gatagcggtt tgactcacgg ggatttccaa gtctccaccc cattgacgtc aatgggagtt 2520
tgttttggca ccaaaatcaa cgggactttc caaaatgtcg taacaactcc gccccattga 2580
cgcaaatggg cggtaggcgt gtacggtggg aggtctatat aagcagagct ctctggctaa 2640
ctagagaacc cactgcttac tggcttatcg aaatgaattc gactcactgt tgggagaccc 2700
aagctggcta gttaagctat cacaagtttg tacaaaaaag caggctggcg cgcccaattc 2760
gccctatagt gagtcgtatt acgtcgcgct cactggccgt cgttttacaa cgtcgtgact 2820
gggaaaaccc tggcgttacc caacttaatc gccttgcagc acatccccct ttcgccagct 2880
ggcgtaatag cgaagaggcc cgcaccgaaa cgcccttccc aacagttgcg cagcctgaat 2940
ggcgaatggg agcgccctgt agcggcgcat taagcgcggc gggtgtggtg gttacgcgca 3000
gcgtgaccgc tacacttgcc agcgccctag cgcccgctcc tttcgctttc ttcccttcct 3060
ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcgggggctc cctttagggt 3120
tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggt gatggttggc 3180
ctgtagtggg ccatagccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 3240
ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc 3300
ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 3360

acaaaaattt aacgcgaatt ttaacaaaat attaacgctt acaatttagg tggcactttt 3420
cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc aaatatgtat 3480
ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag gaagagtatg 3540
agtattcaac atttccgtgt cgcccttatt ccctttttg cggcattttg ccttcctgtt 3600
tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt gggtgcacga 3660
gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa 3720
gaacgttttc caatgatgag cacttttaaa gttctgctat gtggcgcggt attatcccgt 3780
attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt 3840
gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc 3900
agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac aacgatcgga 3960
ggaccgaagg agctaaccgc ttttttgcac aacatggggg atcatgtaac tcgccttgat 4020
cgttgggaac cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct 4080
gtagcaatgg caacaacgtt gcgcaaacta ttaactggcg aactacttac tctagcttcc 4140
cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg 4200
gcccttccgg ctggctggtt tattgctgat aaatctggag ccggtgagcg tggttctcgc 4260
ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg 4320
acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca 4380
ctgattaagc attggtaact gtcagaccaa gtttactcat atatacttta gattgattta 4440
aaacttcatt tttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc 4500
aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa 4560
ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca 4620
ccgctaccag cggtggtttg tttgccggat caagagctac caactctttt tccgaaggta 4680
actggcttca gcagagcgca gataccaaat actgttcttc tagtgtagcc gtagttaggc 4740
caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaat cctgttacca 4800
gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta 4860
ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttggag 4920
cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag cgccacgctt 4980
cccgaaggga gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc 5040
acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac 5100
ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac 5160
gccagcaacg cggccttttt acggttcctg gccttttgct ggccttttgc tcacatggtc 5220
tttcctgcgt tatcccctga ttctgtggat aaccgtatta ccgcctttga gtgagctgat 5280
accgctcgcc gcagccgaac gaccgagcgc agcgagtcag tgagcgagga agcggagagc 5340
gcccaatacg caaaccgcct ctccccgcgc gttggccgat tcattaatgc agctggcacg 5400
acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg agttagctca 5460
ctcattaggc accccaggct ttacacttta tgcttccggc tcgtatgttg tgtggaattg 5520
tgagcggata acaatttcac acaggaaaca gctatgacca tgattacgcc aagcgcgtca 5580
attaaccctc actaaaggga acaaaagctg ttaattaagc gcgccacacg caaacacaaa 5640
tacacagcgg ccttgccacc atggtggatc tacgcacgct cggctacagc cagcagcaac 5700
aggagaagat caaaccgaag gttcgttcga cagtggcgca gcaccacgag gcactggtcg 5760
gccacgggtt tacacacgcg cacatcgttg cgttaagcca acacccggca gcgttaggga 5820
ccgtcgctgt caagtatcag gacatgatcg cagcgttgcc agaggcgaca cacgaagcga 5880
tcgttggcgt cggcaaacag tggtccggcg cacgcgctct ggaggccttg ctcacggtgg 5940
cgggagagtt gagaggtcca ccgttacagt tggacacagg ccaacttctc aagattgcaa 6000
aacgtggcgg cgtgaccgca gtggaggcag tgcatgcatg gcgcaatgca ctgacgggtg 6060
ccccgctcaa cttgaccccg gagcaggtgg tggccatcgc cagccacgat ggcggcaagc 6120
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 6180
cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc 6240
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag gtggtggcca 6300
tcgccagcaa tattggtggc aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc 6360
tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc aatggcggtg 6420
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 6480
tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag gcgctggaga 6540
cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg 6600
tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag gcgctgttgc 6660
cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc gccagcaata 6720
ttggtggcaa gcaggcgctg gagacggtgc aggcgctgtt gccggtgctg tgccaggccc 6780
acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc aagcaggcgc 6840
tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg accccggagc 6900
aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg gtccagcggc 6960
tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg gccatcgcca 7020
gcaatggcgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc 7080
aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc 7140
aggcgctgga gacggtgcag gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc 7200
cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg gagacggtgc 7260
aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag gtggtggcca 7320
tcgccagcca cgatggcggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 7380
tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc cacgatggcg 7440

gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 7500
tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag gcgctggaga 7560
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg 7620
tggccatcgc cagcaatggc ggtggcaagc aggcgctgga gacggtccag cggctgttgc 7680
cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc gccagccacg 7740
atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg tgccaggccc 7800
acggcttgac ccctcagcag gtggtggcca tcgccagcaa tggcggcggc aggccggcgc 7860
tggagagcat tgttgcccag ttatctcgcc ctgatccggg atcctccgga ccgaacagag 7920
gcgttaccaa gcaactagtt aaatccgaac tggaagagaa aaagtccgag ttgcgtcata 7980
aactgaagta cgttccacat gaatacattg aactgatcga gattgcccgt aacagcaccc 8040
aggacagaat cctggaaatg aaagtgatgg agtttttcat gaaagtgtac ggctacagag 8100
gcaaacattt gggcggtagc cgtaaaccgg atggtgccat ctacaccgtt ggcagcccga 8160
tcgattacgg cgtgatcgtt gataccaaag cctactccgg tggctacaac ctgccgattg 8220
gccaggctga cgaaatgcag agatacgtgg aagagaacca gactcgtaac aagcatatca 8280
acccgaacga atggtggaaa gtctacccga gttccgtgac tgagttcaaa tttctgtttg 8340
tgagcggcca tttcaaaggc aactacaaag cccagctgac cagactgaac catatcacca 8400
actgcaacgg tgccgtcttg agcgtggaag aactgttgat tggtggcgaa atgatcaagg 8460
ctggcaccct gactctggag gaagttagac gtaaattcaa caacggtgaa atcaactttg 8520
gatcctcggc cgactgataa gtcgaccacc accaccacca ccactgataa gagctcctcg 8580
agcgctagca cccagctttc ttgtacaaag tggtgatcta gagggcccgc ggttcgaagg 8640
taagcctatc cctaaccctc tcctcggtct cgattctacg cgtaccggtt agtaatgagt 8700
ttaaacgggg gaggctaact gaaacacgga aggagacaat accggaagga acccgcgcta 8760
tgacggcaat aaaaagacag aataaaacgc acgggtgttg ggtcgtttgt tcataaacgc 8820
ggggttcggt cccagggctg gcactctgtc gataccccac cgagacccca ttggggccaa 8880
tacgcccgcg tttcttcctt ttccccaccc cacccccaa gttcgggtga aggcccaggg 8940
ctcgcagcca acgtcggggc ggcaggccct gccatagcag atctgcgcag ctggggctct 9000
aggggtatc cccacgcgcc ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg 9060
cgcagcgtga ccgctacact tgccagcgcc ctagcgcccg ctcctttcgc tttcttccct 9120
tcctttctcg ccacgttcgc cggctttccc cgtcaagctc taaatcgggg catcccttta 9180
gggttccgat ttagtgcttt acggcacctc gaccccaaaa aacttgatta gggtgatggt 9240
tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg 9300
ttctttaata gtggactctt gttccaaact ggaacaacac tcaaccctat ctcggtctat 9360
tcttttgatt tataagggat tttggggatt tcggcctatt ggttaaaaaa tgagctgatt 9420
taacaaaaat ttaacgcgaa ttaattctgt ggaatgtgtg tcagttaggg tgtggaaagt 9480
ccccaggctc cccagcaggc agaagtatgc aaagcatgca tctcaattag tcagcaacca 9540
ggtgtggaaa gtccccaggc tccccagcag gcagaagtat gcaaagcatg catctcaatt 9600
agtcagcaac catagtcccg cccctaactc cgcccatccc gcccctaact ccgcccagtt 9660
ccgcccattc tccgccccat ggctgactaa ttttttttat ttatgcagag gccgaggccg 9720
cctctgcctc tgagctattc cagaagtagt gaggaggctt ttttggaggc ctaggctttt 9780
gcaaaaagct cccgggagct tgtatatcca ttttcggatc tgatcagcac gtgttgacaa 9840
ttaatcatcg gcatagtata tcggcatagt ataatacgac aaggtgagga actaaaccat 9900
ggccaagcct ttgtctcaag aagaatccac cctcattgaa agagcaacgg ctacaatcaa 9960
cagcatcccc atctctgaag actacagcgt cgccagcgca gctctctcta gcgacggccg 10020
catcttcact ggtgtcaatg tatatcattt tactggggga ccttgtgcag aactcgtggt 10080
gctgggcact gctgctgctg cggcagctgg caacctgact tgtatcgtcg cgatcggaaa 10140
tgagaacagg ggcatcttga gcccctgcgg acggtgccga caggtgcttc tcgatctgca 10200
tcctgggatc aaagccatag tgaaggacag tgatggacag ccgacggcag ttgggattcg 10260
tgaattgctg ccctctggtt atgtgtggga gggctaagca cttcgtggcc gaggagcagg 10320
actgacacgt gctacgagat ttcgattcca ccgccgcctt ctatgaaagg ttgggcttcg 10380
gaatcgtttt ccgggacgcc ggctggatga tcctccagcg cggggatctc atgctggagt 10440
tcttcgccca ccccaacttg tttattgcag cttataatgg ttacaaataa agcaatagca 10500
tcacaaattt cacaaataaa gcatttttt cactgcattc tagttgtggt ttgtccaaac 10560
tcatcaatgt atcttatcat gtctgtatac cgtcgacctc tagctagagc ttggcgtaat 10620
catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac 10680
gagccggaag cataaagtgt aaagcctggg gtgcctaatg agtgagctaa ctcacattaa 10740
ttgcgttgcg ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat 10800
gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc 10860
tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 10920
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaac 10967

<210> 359

<211> 12449
<212> DNA
<213> Artificial sequence

<220>

<223> pCLS14971

<400> 359

```
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt     60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg    120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc    180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc    240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc    300
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac    360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt    420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct    480
aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc    540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtttt    600
tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc    660
ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg    720
agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca    780
atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca    840
cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag    900
ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac    960
ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc   1020
agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct   1080
agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc   1140
gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg   1200
cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc   1260
gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat   1320
tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag   1380
tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat   1440
aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg   1500
cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca   1560
cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga   1620
aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc   1680
ttccttttc aatattattg aagcatttat cagggttatt gtctcatgag cggatacata   1740
tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg   1800
ccacctgacg tcgacggatc gggagatctc ccgatcccct atggtgcact ctcagtacaa   1860
tctgctctga tgccgcatag ttaagccagt atctgctccc tgcttgtgtg ttggaggtcg   1920
ctgagtagtg cgcgagcaaa atttaagcta caacaaggca aggcttgacc gacaattgca   1980
tgaagaatct gcttagggtt aggcgttttg cgctgcttcg cgatgtacgg gccagatata   2040
cgcgttgaca ttgattattg actagttatt aatagtaatc aattacgggg tcattagttc   2100
atagcccata tatggagttc cgcgttacat aacttacggt aaatggcccg cctggctgac   2160
cgcccaacga cccccgccca ttgacgtcaa taatgacgta tgttcccata gtaacgccaa   2220
tagggacttt ccattgacgt caatgggtgg agtatttacg gtaaactgcc cacttggcag   2280
tacatcaagt gtatcatatg ccaagtacgc cccctattga cgtcaatgac ggtaaatggc   2340
ccgcctggca ttatgcccag tacatgacct tatgggactt tcctacttgg cagtacatct   2400
acgtattagt catcgctatt accatggtga tgcggttttg gcagtacatc aatgggcgtg   2460
gatagcggtt tgactcacgg ggatttccaa gtctccaccc cattgacgtc aatgggagtt   2520
tgttttggca ccaaaatcaa cgggactttc caaaatgtcg taacaactcc gccccattga   2580
cgcaaatggg cggtaggcgt gtacggtggg aggtctatat aagcagagct ctctggctaa   2640
ctagagaacc cactgcttac tggcttatcg aaatgaattc gactcactgt tgggagaccc   2700
aagctggcta gttaagctat cacaagtttg tacaaaaaag caggctggcg cgcccaattc   2760
gccctatagt gagtcgtatt acgtcgcgct cactggccgt cgttttacaa cgtcgtgact   2820
gggaaaaccc tggcgttacc caacttaatc gccttgcagc acatccccct ttcgccagct   2880
ggcgtaatag cgaagaggcc cgcaccgaaa cgcccttccc aacagttgcg cagcctgaat   2940
ggcgaatggg agcgccctgt agcggcgcat taagcgcggc gggtgtggtg gttacgcgca   3000
gcgtgaccgc tacacttgcc agcgccctag cgcccgctcc tttcgctttc ttcccttcct   3060
ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcgggggctc cctttagggt   3120
tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggt gatggttggc   3180
ctgtagtggg ccatagccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt   3240
ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc   3300
ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta   3360
acaaaaattt aacgcgaatt ttaacaaaat attaacgctt acaatttagg tggcactttt   3420
cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc aaatatgtat   3480
ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag gaagagtatg   3540
agtattcaac atttccgtgt cgcccttatt cccttttttg cggcattttg ccttcctgtt   3600
tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt gggtgcacga   3660
gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa   3720
gaacgttttc caatgatgag cacttttaaa gttctgctat gtggcgcggt attatcccgt   3780
attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt   3840
gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc   3900
```

agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac aacgatcgga 3960
ggaccgaagg agctaaccgc ttttttgcac aacatggggg atcatgtaac tcgccttgat 4020
cgttgggaac cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct 4080
gtagcaatgg caacaacgtt gcgcaaacta ttaactggcg aactacttac tctagcttcc 4140
cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg 4200
gcccttccgg ctggctggtt tattgctgat aaatctggag ccggtgagcg tggttctcgc 4260
ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg 4320
acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca 4380
ctgattaagc attggtaact gtcagaccaa gtttactcat atatacttta gattgattta 4440
aaacttcatt tttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc 4500
aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa 4560
ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca 4620
ccgctaccag cggtggtttg tttgccggat caagagctac caactctttt tccgaaggta 4680
actggcttca gcagagcgca gataccaaat actgttcttc tagtgtagcc gtagttaggc 4740
caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaat cctgttacca 4800
gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta 4860
ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttggag 4920
cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag cgccacgctt 4980
cccgaaggga gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc 5040
acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac 5100
ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac 5160
gccagcaacg cggccttttt acggttcctg gccttttgct ggccttttgc tcacatggtc 5220
tttcctgcgt tatcccctga ttctgtggat aaccgtatta ccgcctttga gtgagctgat 5280
accgctcgcc gcagccgaac gaccgagcgc agcgagtcag tgagcgagga agcggagagc 5340
gcccaatacg caaaccgcct ctccccgcgc gttggccgat tcattaatgc agctggcacg 5400
acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg agttagctca 5460
ctcattaggc accccaggct ttacacttta tgcttccggc tcgtatgttg tgtggaattg 5520
tgagcggata acaatttcac acaggaaaca gctatgacca tgattacgcc aagcgcgtca 5580
attaaccctc actaaaggga acaaaagctg ttaattaagc gcgccacacg caaacacaaa 5640
tacacagcgg ccttgccacc atgggttccg aggcaccccg ggccgagacc tttgtcttcc 5700
tggacctgga agccactggg ctccccagtg tggagcccga gattgccgag ctgtccctct 5760
ttgctgtcca ccgctcctcc ctggagaacc cggagcacga cgagtctggt gccctagtat 5820
tgccccgggt cctggacaag ctcacgctgt gcatgtgccc ggagcgcccc ttcactgcca 5880
aggccagcga gatcaccggc ctgagcagtg agggcctggc gcgatgccgg aaggctggct 5940
ttgatggcgc cgtggtgcgg acgctgcagg ccttcctgag ccgccaggca gggcccatct 6000
gccttgtggc ccacaatggc tttgattatg atttcccccc gctgtgtgcc gagctgcggc 6060
gcctgggtgc ccgcctgccc cgggacactg tctgcctgga cacgctgccg gccctgcggg 6120
gcctggaccg cgcccacagc cacggcaccc gggcccgggg ccgccagggt tacagcctcg 6180
gcagcctctt ccaccgctac ttccgggcag agccaagcgc agcccactca gccgagggcg 6240
acgtgcacac cctgctcctg atcttcctgc accgcgccgc agagctgctc gcctgggccg 6300
atgagcaggc ccgtgggtgg gcccacatcg agcccatgta cttgccgcct gatgacccca 6360
gcctggaggc gactcctcca cagaccggtc tggatgttcc ttactccgag gcaccccggg 6420
ccgagacctt tgtcttcctg gacctggaag ccactgggct ccccagtgtg gagcccgaga 6480
ttgccgagct gtccctcttt gctgtccacc gctcctccct ggagaacccg gagcacgacg 6540
agtctggtgc cctagtattg ccccgggtcc tggacaagct cacgctgtgc atgtgcccgg 6600
agcgcccctt cactgccaag gccagcgaga tcaccggcct gagcagtgag ggcctggcgc 6660
gatgccggaa ggctggcttt gatggcgccg tggtgcggac gctgcaggcc ttcctgagcc 6720
gccaggcagg gcccatctgc cttgtggccc acaatggctt tgattatgat ttccccctgc 6780
tgtgtgccga gctgcggcgc ctgggtgccc gcctgccccg ggacactgtc tgcctggaca 6840
cgctgccggc cctgcggggc ctggaccgcg cccacagcca cggcacccgg gcccggggcc 6900
gccagggtta cagcctcggc agcctcttcc accgctactt ccgggcagag ccaagcgcag 6960
cccactcagc cgagggcgac gtgcacaccc tgctcctgat cttcctgcac cgcgccgcag 7020
agctgctcgc ctgggccgat gagcaggccc gtgggtgggc ccacatcgag cccatgtact 7080
tgccgcctga tgaccccagc ctggaggcgg gaggtggagg ttctggaggt ggaggatcct 7140
ccggagtgga tctacgcacg ctcggctaca gccagcagca acaggagaag atcaaaccga 7200
aggttcgttc gacagtggcg cagcaccacg aggcactggt cggccacggg tttacacacg 7260
cgcacatcgt tgcgttaagc caacacccgg cagcgttagg gaccgtcgct gtcaagtatc 7320
aggacatgat cgcagcgttg ccagaggcga cacacgaagc gatcgttggc gtcggcaaac 7380
agtggtccgg cgcacgcgct ctggaggcct tgctcacggt ggcgggagag ttgagaggtc 7440
caccgttaca gttggacaca ggccaacttc tcaagattgc aaaacgtggc ggcgtgaccg 7500
cagtggaggc agtgcatgca tggcgcaatg cactgacggg tgccccgctc aacttgaccc 7560
cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg gagacggtcc 7620
agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag gtggtggcca 7680
tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg ttgccggtgc 7740
tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc aatattggtg 7800
gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag gcccacggct 7860
tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag gcgctggaga 7920
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg 7980

tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag gcgctgttgc 8040
cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc gccagcaata 8100
ttggtggcaa gcaggcgctg gagacggtgc aggcgctgtt gccggtgctg tgccaggccc 8160
acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc aagcaggcgc 8220
tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg accccggagc 8280
aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg gtccagcggc 8340
tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg gccatcgcca 8400
gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc 8460
aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc ggtggcaagc 8520
aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc 8580
cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg gagacggtgc 8640
aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag gtggtggcca 8700
tcgccagcaa tattggtggc aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc 8760
tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc cacgatggcg 8820
gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag gcccacggct 8880
tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag gcgctggaga 8940
cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg 9000
tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag cggctgttgc 9060
cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc gccagcaatg 9120
gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg tgccaggccc 9180
acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc aagcaggcgc 9240
tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg acccctcagc 9300
aggtggtggc catcgccagc aatggcggcg gcaggccggc gctggagagc attgttgccc 9360
agttatctcg ccctgatccg ggatcctccg gaccgaacag aggcgttacc aagcaactag 9420
ttaaatccga actggaagag aaaaagtccg agttgcgtca taaactgaag tacgttccac 9480
atgaatacat tgaactgatc gagattgccc gtaacagcac ccaggacaga atcctggaaa 9540
tgaaagtgat ggagtttttc atgaaagtgt acggctacag aggcaaacat ttgggcggta 9600
gccgtaaacc ggatggtgcc atctacaccg ttggcagccc gatcgattac ggcgtgatcg 9660
ttgataccaa agcctactcc ggtggctaca acctgccgat tggccaggct gacgaaatgc 9720
agagatacgt ggaagagaac cagactcgta acaagcatat caacccgaac gaatggtgga 9780
aagtctaccc gagttccgtg actgagttca aatttctgtt tgtgagcggc catttcaaag 9840
gcaactacaa agcccagctg accagactga accatatcac caactgcaac ggtgccgtct 9900
tgagcgtgga agaactgttg attggtggcg aaatgatcaa ggctggcacc ctgactctgg 9960
aggaagttag acgtaaattc aacaacggtg aaatcaactt tggatcctcg gccgactgat 10020
aagtcgacca ccaccaccac caccactgat aagagctcct cgagcgctag cacccagctt 10080
tcttgtacaa agtggtgatc tagagggccc gcggttcgaa ggtaagccta tccctaaccc 10140
tctcctcggt ctcgattcta cgcgtaccgg ttagtaatga gtttaaacgg gggaggctaa 10200
ctgaaacacg gaaggagaca ataccggaag gaacccgcgc tatgacggca ataaaaagac 10260
agaataaaac gcacgggtgt tgggtcgttt gttcataaac gcggggttcg gtcccagggc 10320
tggcactctg tcgatacccc accgagaccc cattggggcc aatacgcccg cgtttcttcc 10380
ttttccccac cccacccccc aagttcgggt gaaggcccag ggctcgcagc caacgtcggg 10440
gcggcaggcc ctgccatagc agatctgcgc agctggggct ctaggggtga tccccacgcg 10500
ccctgtagcg gcgcattaag cgcggcgggt gtggtggtta cgcgcagcgt gaccgctaca 10560
cttgccagcg ccctagcgcc cgctcctttc gctttcttcc cttcctttct cgccacgttc 10620
gccggctttc cccgtcaagc tctaaatcgg ggcatcccct tagggttccg atttagtgct 10680
ttacggcacc tcgaccccaa aaaacttgat tagggtgatg gttcacgtag tgggccatcg 10740
ccctgataga cggtttttcg ccctttgacg ttggagtcca cgttctttaa tagtggactc 10800
ttgttccaaa ctggaacaac actcaaccct atctcggtct attcttttga tttataaggg 10860
attttgggga tttcggccta ttggttaaaa aatgagctga tttaacaaaa atttaacgcg 10920
aattaattct gtggaatgtg tgtcagttag ggtgtggaaa gtccccaggc tccccagcag 10980
gcagaagtat gcaaagcatg catctcaatt agtcagcaac caggtgtgga aagtccccag 11040
gctccccagc aggcagaagt atgcaaagca tgcatctcaa ttagtcagca accatagtcc 11100
cgcccctaac tccgcccatc ccgcccctaa ctccgcccag ttccgcccat tctccgcccc 11160
atggctgact aatttttttt atttatgcag aggccgaggc cgcctctgcc tctgagctat 11220
tccagaagta gtgaggaggc ttttttggag gcctaggctt ttgcaaaaag ctcccgggag 11280
cttgtatatc cattttcgga tctgatcagc acgtgttgac aattaatcat cggcatagta 11340
tatcggcata gtataatacg acaaggtgag gaactaaacc atggccaagc ctttgtctca 11400
agaagaatcc accctcattg aaagagcaac ggctacaatc aacagcatcc ccatctctga 11460
agactacagc gtcgccagcg cagctctctc tagcgacggc cgcatcttca ctggtgtcaa 11520
tgtatatcat tttactgggg gaccttgtgc agaactcgtg gtgctgggca ctgctgctgc 11580
tgcggcagct ggcaacctga cttgtatcgt cgcgatcgga aatgagaaca ggggcatctt 11640
gagcccctgc ggacggtgcc gacaggtgct tctcgatctg catcctggga tcaaagccat 11700
agtgaaggac agtgatggac agccgacggc agttgggatt cgtgaattgc tgccctctgg 11760
ttatgtgtgg gagggctaag cacttcgtgg ccgaggagca ggactgacac gtgctacgag 11820
atttcgattc caccgccgcc ttctatgaaa ggttgggctt cggaatcgtt ttccgggacg 11880
ccggctggat gatcctccag cgcggggatc tcatgctgga gttcttcgcc caccccaact 11940
tgtttattgc agcttataat ggttacaaat aaagcaatag catcacaaat ttcacaaata 12000
aagcattttt ttcactgcat tctagttgtg gtttgtccaa actcatcaat gtatcttatc 12060

atgtctgtat accgtcgacc tctagctaga gcttggcgta atcatggtca tagctgtttc 12120
ctgtgtgaaa ttgttatccg ctcacaattc cacacaacat acgagccgga agcataaagt 12180
gtaaagcctg gggtgcctaa tgagtgagct aactcacatt aattgcgttg cgctcactgc 12240
ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg 12300
ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc gctcactgac tcgctgcgct 12360
cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca 12420
cagaatcagg ggataacgca ggaaagaac 12449

<210> 360

<211> 7997
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15800

<400> 360
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aaggaggaaa cattcatgga 420
actgaagaac tccatcagcg attacaccga agccgagttc gtgcagctgt tgaaagaaat 480
tgagaaagag aacgttgctg ccaccgatga cgtgctggat gttttgctgg aacattttgt 540
gaagatcact gaacatccag atggcactga cctgatctac tacccgagcg ataaccgtga 600
tgacagccca gaaggcattg ttaaagagat caaagaatgg agagctgcca acggcaagcc 660
aggcttcaaa cagggttgac ggccaactga tttaattaag cgcgccacac gcaaacacaa 720
atacacagcg gccttgccac catggtggat ctacgcacgc tcggctacag ccagcagcaa 780
caggagaaga tcaaaccgaa ggttcgttcg acagtggcgc agcaccacga ggcactggtc 840
ggccacgggt ttacacacgc gcacatcgtt gcgttaagcc aacacccggc agcgttaggg 900
accgtcgctg tcaagtatca ggacatgatc gcagcgttgc cagaggcgac acacgaagcg 960
atcgttggcg tcggcaaaca gtggtccggc gcacgcgctc tggaggcctt gctcacggtg 1020
gcgggagagt tgagaggtcc accgttacag ttggacacag gccaacttct caagattgca 1080
aaacgtggcg gcgtgaccgc agtggaggca gtgcatgcat ggcgcaatgc actgacgggt 1140
gccccgctca acttgacccc ggagcaggtg gtggccatcg ccagccacga tggcggcaag 1200
caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc 1260
ccccagcagg tggtggccat cgccagcaat ggcggtggca agcaggcgct ggagacggtc 1320
cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc 1380
atcgccagca atattggtgg caagcaggcg ctggagacgg tgcaggcgct gttgccggtg 1440
ctgtgccagg cccacggctt gacccccag caggtggtgg ccatcgccag caatggcggt 1500
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc 1560
ttgaccccgg agcaggtggt ggccatcgcc agcaatattg gtggcaagca ggcgctggag 1620
acggtgcagg cgctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg 1680
gtggccatcg ccagcaatat tggtggcaag caggcgctgg agacggtgca ggcgctgttg 1740
ccggtgctgt gccaggccca cggcttgacc ccggagcagg tggtggccat cgccagcaat 1800
attggtggca agcaggcgct ggagacggtg caggcgctgt tgccggtgct gtgccaggcc 1860
cacggcttga ccccggagca ggtggtggcc atcgccagcc acgatggcgg caagcaggcg 1920
ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt gaccccggag 1980
caggtggtgg ccatcgccag ccacgatggc ggcaagcagg cgctggagac ggtccagcgg 2040
ctgttgccgg tgctgtgcca ggcccacggc ttgacccccc agcaggtggt ggccatcgcc 2100
agcaatggcg gtggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc 2160
caggcccacg gcttgacccc ggagcaggtg gtggccatcg ccagcaatat tggtggcaag 2220
caggcgctgg agacggtgca ggcgctgttg ccggtgctgt gccaggccca cggcttgacc 2280
ccggagcagg tggtggccat cgccagcaat attggtggca agcaggcgct ggagacggtg 2340
caggcgctgt tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc 2400
atcgccagcc acgatggcgg caagcaggcg ctggagacgg tccagcggct gttgccggtg 2460
ctgtgccagg cccacggctt gaccccggag caggtggtgg ccatcgccag ccacgatggc 2520
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc 2580
ttgaccccgg agcaggtggt ggccatcgcc agccacgatg gcggcaagca ggcgctggag 2640
acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc ccagcaggtg 2700
gtggccatcg ccagcaatgg cggtggcaag caggcgctgg agacggtcca gcggctgttg 2760
ccggtgctgt gccaggccca cggcttgacc ccggagcagg tggtggccat cgccagccac 2820
gatggcggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc 2880
cacggcttga cccctcagca ggtggtggcc atcgccagca atggcggcgg caggccggcg 2940

ctggagagca ttgttgccca gttatctcgc cctgatccgg gatcctccgg atctaaacgc 3000
aacaaaccag gtaaggccac tggcaaaggt aaaccagtga acaataagtg gctgaacaac 3060
gctggcaaag acttgggttc tccagtgccg gatcgcattg ccaacaagct gagagacaaa 3120
gaattcaaga gttttgatga ctttcgcgca aaattctggg aagaagtgtc taaagatcca 3180
gaactgtcta aacagtttag tcgcaacaac aatgacagaa tgaaagttgg taaagctccg 3240
aagacccgca cccaagacgt gtctggcaaa cgcaccagtt ttgaattgca tcatgaaaaa 3300
ccaatctctc agaatggtgg cgtttacgac atggacaaca tctctgtggt tactccgaaa 3360
agacatatcg atattcatcg cggtaaagga tccaccggtg gatcctccgg agtggatcta 3420
cgcacgctcg gctacagcca gcagcaacag gagaagatca aaccgaaggt tcgttcgaca 3480
gtggcgcagc accacgaggc actggtcggc cacgggttta cacacgcgca catcgttgcg 3540
ttaagccaac acccggcagc gttagggacc gtcgctgtca agtatcagga catgatcgca 3600
gcgttgccag aggcgacaca cgaagcgatc gttggcgtcg gcaaacagtg gtccggcgca 3660
cgcgctctgg aggccttgct cacggtggcg ggagagttga gaggtccacc gttacagttg 3720
gacacaggcc aacttctcaa gattgcaaaa cgtggcggcg tgaccgcagt ggaggcagtg 3780
catgcatggc gcaatgcact gacgggtgcc ccgctcaact tgacccccca gcaggtggtg 3840
gccatcgcca gcaataatgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg 3900
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 3960
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 4020
ggcttgaccc cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg 4080
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 4140
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 4200
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 4260
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 4320
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 4380
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc 4440
cagcaggtgg tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag 4500
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 4560
gccagcaata atggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 4620
tgccaggccc acggcttgac cccccagcag gtggtggcca tcgccagcaa tggcggtggc 4680
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 4740
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 4800
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 4860
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 4920
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat 4980
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 5040
ggcttgaccc cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg 5100
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 5160
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 5220
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 5280
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 5340
gcccacggct tgacccctca gcaggtggtg gccatcgcca gcaatggcgg cggcaggccg 5400
gcgctggaga gcattgttgc ccagttatct cgccctgatc cgtcggccga ctgataagtc 5460
gaccaccacc accaccacca ctgataagag ctcctcgagc ctgcagcagc tgaagctttg 5520
gacttctggc gcgcccaatt cgccctatag tgagtcgtat tacgtcgcgc tcactggccg 5580
tcgttttaca acgtcgtgac tgggaaaacc ctggcgttac ccaacttaat cgccttgcag 5640
cacatccccc tttcgccagc tggcgtaata gcgaagaggc ccgcaccgaa acgcccttcc 5700
caacagttgc gcagcctgaa tggcgaatgg gagcgccctg tagcggcgca ttaagcgcgg 5760
cgggtgtggt ggttacgcgc agcgtgaccg ctacacttgc cagcgcccta gcgcccgctc 5820
ctttcgcttt cttcccttcc tttctcgcca cgttcgccgg ctttccccgt caagctctaa 5880
atcgggggct ccctttaggg ttccgattta gtgctttacg gcacctcgac cccaaaaaac 5940
ttgattaggg tgatggttgg cctgtagtgg gccatagccc tgatagacgg tttttcgccc 6000
tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact 6060
caaccctatc tcggtctatt cttttgattt ataagggatt ttgccgattt cggcctattg 6120
gttaaaaaat gagctgattt aacaaaaatt taacgcgaat tttaacaaaa tattaacgct 6180
tacaatttag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc 6240
taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa 6300
tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt 6360
gcggcatttt gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct 6420
gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc 6480
cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta 6540
tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac 6600
tattctcaga atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc 6660
atgacagtaa gagaattatg cagtgctgcc ataaccatga gtgataacac tgcggccaac 6720
ttacttctga caacgatcgg aggaccgaag gagctaaccg ctttttgca caacatgggg 6780
gatcatgtaa ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac 6840
gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc 6900
gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt 6960
gcaggaccac ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga 7020

gccggtgagc gtggttctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc 7080
cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag 7140
atcgctgaga taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca 7200
tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc 7260
ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca 7320
gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc 7380
tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta 7440
ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt 7500
ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc 7560
gctctgctaa tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg 7620
ttggactcaa gacgatagtt accggataag gcgcagcggt cgggctgaac gggggttcg 7680
tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag 7740
ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc 7800
agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat 7860
agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg 7920
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 7980
tggccttttg ctcacat 7997

<210> 361

<211> 7997
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15801

<400> 361
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aaggaggaaa cattcatgga 420
actgaagaac tccatcagcg attacaccga agccgagttc gtgcagctgt tgaaagaaat 480
tgagaaagag aacgttgctg ccaccgatga cgtgctggat gttttgctgg aacattttgt 540
gaagatcact gaacatccag atggcactga cctgatctac tacccgagcg ataaccgtga 600
tgacagccca gaaggcattg ttaaagagat caaagaatgg agagctgcca acggcaagcc 660
aggcttcaaa cagggttgac ggccaactga tttaattaag cgcgccacac gcaaacacaa 720
atacacagcg gccttgccac catggtggat ctacgcacgc tcggctacag ccagcagcaa 780
caggagaaga tcaaaccgaa ggttcgttcg acagtggcgc agcaccacga ggcactggtc 840
ggccacgggt ttacacacgc gcacatcgtt gcgttaagcc aacacccggc agcgttaggg 900
accgtcgctg tcaagtatca ggacatgatc gcagcgttgc cagaggcgac acacgaagcg 960
atcgttggcg tcggcaaaca gtggtccggc gcacgcgctc tggaggcctt gctcacggtg 1020
gcgggagagt tgagaggtcc accgttacag ttggacacag gccaacttct caagattgca 1080
aaacgtggcg gcgtgaccgc agtggaggca gtgcatgcat ggcgcaatgc actgacgggt 1140
gccccgctca acttgacccc ccagcaggtg gtggccatcg ccagcaataa tggtggcaag 1200
caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc 1260
ccccagcagg tggtggccat cgccagcaat ggcggtggca agcaggcgct ggagacggtc 1320
cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccccagca ggtggtggcc 1380
atcgccagca atggcggtgg caagcaggcg ctggagacgg tccagcggct gttgccggtg 1440
ctgtgccagg cccacggctt gacccccag caggtggtgg ccatcgccag caatggcggt 1500
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc 1560
ttgaccccgg agcaggtggt ggccatcgcc agcaatattg gtggcaagca ggcgctggag 1620
acggtgcagg cgctgttgcc ggtgctgtgc caggcccacg gcttgacccc ccagcaggtg 1680
gtggccatcg ccagcaatgg cggtggcaag caggcgctgg agacggtcca gcggctgttg 1740
ccggtgctgt gccaggccca cggcttgacc ccccagcagg tggtggccat cgccagcaat 1800
aatggtggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc 1860
cacggcttga ccccccagca ggtggtggcc atcgccagca ataatggtgg caagcaggcg 1920
ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt gacccccag 1980
caggtggtgg ccatcgccag caatggcggt ggcaagcagg cgctggagac ggtccagcgg 2040
ctgttgccgg tgctgtgcca ggcccacggc ttgacccccc agcaggtggt ggccatcgcc 2100
agcaatggcg gtggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc 2160
caggcccacg gcttgacccc ggagcaggtg gtggccatcg ccagcaatat tggtggcaag 2220
caggcgctgg agacggtgca ggcgctgttg ccggtgctgt gccaggccca cggcttgacc 2280
ccggagcagg tggtggccat cgccagccac gatggcggca agcaggcgct ggagacggtc 2340

```
cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccccagca ggtggtggcc   2400
atcgccagca atggcggtgg caagcaggcg ctggagacgg tccagcggct gttgccggtg   2460
ctgtgccagg cccacggctt gacccccag caggtggtgg ccatcgccag caatggcggt   2520
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc   2580
ttgaccccgg agcaggtggt ggccatcgcc agcaatattg gtggcaagca ggcgctggag   2640
acggtgcagg cgctgttgcc ggtgctgtgc caggcccacg gcttgacccc tcagcaggtg   2700
gtggccatcg ccagcaatgg cggcggcagg ccggcgctgg agagcattgt tgcccagtta   2760
tctcgccctg atccgggatc ctccggatct aaacgcaaca aaccaggtaa ggccactggc   2820
aaaggtaaac cagtgaacaa taagtggctg aacaacgctg gcaaagactt gggttctcca   2880
gtgccggatc gcattgccaa caagctgaga gacaaagaat tcaagagttt tgatgacttt   2940
cgcgcaaaat tctgggaaga agtgtctaaa gatccagaac tgtctaaaca gtttagtcgc   3000
aacaacaatg acagaatgaa agttggtaaa gctccgaaga cccgcaccca agacgtgtct   3060
ggcaaacgca ccagttttga attgcatcat gaaaaaccaa tctctcagaa tggtggcgtt   3120
tacgacatgg acaacatctc tgtggttact ccgaaaagac atatcgatat tcatcgcggt   3180
aaaggatcca ccggtggatc ctccggagtg gatctacgca cgctcggcta cagccagcag   3240
caacaggaga agatcaaacc gaaggttcgt tcgacagtgg cgcagcacca cgaggcactg   3300
gtcggccacg ggtttacaca cgcgcacatc gttgcgttaa gccaacaccc ggcagcgtta   3360
gggaccgtcg ctgtcaagta tcaggacatg atcgcagcgt tgccagaggc gacacacgaa   3420
gcgatcgttg gcgtcggcaa acagtggtcc ggcgcacgcg ctctggaggc cttgctcacg   3480
gtggcgggag agttgagagg tccaccgtta cagttggaca caggccaact tctcaagatt   3540
gcaaaacgtg gcggcgtgac cgcagtggag gcagtgcatg catggcgcaa tgcactgacg   3600
ggtgccccgc tcaacttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc   3660
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg   3720
acccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg   3780
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg   3840
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg   3900
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc   3960
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac   4020
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg   4080
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccggagcag   4140
gtggtggcca tcgccagcaa tattggtggc aagcaggcgc tggagacggt gcaggcgctg   4200
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc   4260
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag   4320
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag   4380
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg   4440
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag   4500
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc   4560
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg   4620
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc   4680
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg   4740
accccggagc aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg   4800
gtgcaggcgc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg   4860
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg   4920
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat   4980
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac   5040
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg   5100
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag   5160
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg   5220
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc   5280
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag   5340
gcccacggct tgacccctca gcaggtggtg gccatcgcca gcaatggcgg cggcaggccg   5400
gcgctggaga gcattgttgc ccagttatct cgccctgatc cgtcggccga ctgataagtc   5460
gaccaccacc accaccacca ctgataagag ctcctcgagc ctgcagcagc tgaagctttg   5520
gacttctggc gcgcccaatt cgccctatag tgagtcgtat tacgtcgcgc tcactggccg   5580
tcgttttaca acgtcgtgac tgggaaaacc ctggcgttac ccaacttaat cgccttgcag   5640
cacatccccc tttcgccagc tggcgtaata gcgaagaggc ccgcaccgaa acgcccttcc   5700
caacagttgc gcagcctgaa tggcgaatgg gagcgccctg tagcggcgca ttaagcgcgg   5760
cgggtgtggt ggttacgcgc agcgtgaccg ctacacttgc cagcgcccta gcgcccgctc   5820
ctttcgcttt cttcccttcc tttctcgcca cgttcgccgg ctttccccgt caagctctaa   5880
atcgggggct ccctttaggg ttccgattta gtgctttacg gcacctcgac cccaaaaaac   5940
ttgattaggg tgatggttgg cctgtagtgg gccatagccc tgatagacgg tttttcgccc   6000
tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact   6060
caaccctatc tcggtctatt cttttgattt ataagggatt ttgccgattt cggcctattg   6120
gttaaaaaat gagctgattt aacaaaaatt taacgcgaat tttaacaaaa tattaacgct   6180
tacaatttag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc   6240
taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa   6300
tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt   6360
gcggcatttt gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct   6420
```

gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc 6480
cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta 6540
tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac 6600
tattctcaga atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc 6660
atgacagtaa gagaattatg cagtgctgcc ataaccatga gtgataacac tgcggccaac 6720
ttacttctga caacgatcgg aggaccgaag gagctaaccg ctttttgca caacatgggg 6780
gatcatgtaa ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac 6840
gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc 6900
gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt 6960
gcaggaccac ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga 7020
gccggtgagc gtggttctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc 7080
cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag 7140
atcgctgaga taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca 7200
tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc 7260
ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca 7320
gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc 7380
tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta 7440
ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt 7500
ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc 7560
gctctgctaa tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg 7620
ttggactcaa gacgatagtt accggataag gcgcagcggt cgggctgaac gggggttcg 7680
tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag 7740
ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc 7800
agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat 7860
agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg 7920
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 7980
tggccttttg ctcacat 7997

<210> 362

<211> 13011
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS12106

<400> 362
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740

atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg tggatctacg cacgctcggc 2460
tacagccagc agcaacagga gaagatcaaa ccgaaggttc gttcgacagt ggcgcagcac 2520
cacgaggcac tggtcggcca cgggtttaca cacgcgcaca tcgttgcgtt aagccaacac 2580
ccggcagcgt tagggaccgt cgctgtcaag tatcaggaca tgatcgcagc gttgccagag 2640
gcgacacacg aagcgatcgt tggcgtcggc aaacagtggt ccggcgcacg cgctctggag 2700
gccttgctca cggtggcggg agagttgaga ggtccaccgt tacagttgga cacaggccaa 2760
cttctcaaga ttgcaaaacg tggcggcgtg accgcagtgg aggcagtgca tgcatggcgc 2820
aatgcactga cgggtgcccc gctcaacttg accccggagc aggtggtggc catcgccagc 2880
cacgatggcg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2940
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 3000
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 3060
gagcaggtgg tggccatcgc cagcaatatt ggtggcaagc aggcgctgga gacggtgcag 3120
gcgctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 3180
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3240
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 3300
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 3360
accccggagc aggtggtggc catcgccagc aatattggtg gcaagcaggc gctggagacg 3420
gtgcaggcgc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 3480
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg 3540
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat 3600
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 3660
ggcttgaccc cggagcaggt ggtggccatc gccagccacg atggcggcaa gcaggcgctg 3720
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 3780
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 3840
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 3900
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 3960
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gcaatattgg tggcaagcag 4020
gcgctggaga cggtgcaggc gctgttgccg gtgctgtgcc aggcccacgg cttgaccccg 4080
gagcaggtgg tggccatcgc cagccacgat ggcggcaagc aggcgctgga gacggtccag 4140
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cggagcaggt ggtggccatc 4200
gccagccacg atggcggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 4260
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcca cgatggcggc 4320
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 4380
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 4440
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg 4500
gccatcgcca gccacgatgg cggcaagcag gcgctggaga cggtccagcg gctgttgccg 4560
gtgctgtgcc aggcccacgg cttgacccct cagcaggtgg tggccatcgc cagcaatggc 4620
ggcggcaggc cggcgctgga gagcattgtt gcccagttat ctcgccctga tccgggatcc 4680
tccggatcta aacgcaacaa accaggtaag gccactggca aaggtaaacc agtgaacaat 4740
aagtggctga acaacgctgg caaagacttg ggttctccag tgccggatcg cattgccaac 4800
aagctgagag acaaagaatt caagagtttt gatgactttc gcgcaaaatt ctgggaagaa 4860
gtgtctaaag atccagaact gtctaaacag tttagtcgca acaacaatga cagaatgaaa 4920
gttggtaaag ctccgaagac ccgcacccaa gacgtgtctg gcaaacgcac cagttttgaa 4980
ttgcatcatg aaaaaccaat ctctcagaat ggtggcgttt acgacatgga caacatctct 5040
gtggttactc cgaaaagaca tatcgatatt catcgcggta aaggatccac cggtggatcc 5100
tccggagtgg atctacgcac gctcggctac agccagcagc aacaggagaa gatcaaaccg 5160
aaggttcgtt cgacagtggc gcagcaccac gaggcactgg tcggccacgg gtttacacac 5220
gcgcacatcg ttgcgttaag ccaacacccg gcagcgttag ggaccgtcgc tgtcaagtat 5280
caggacatga tcgcagcgtt gccagaggcg acacacgaag cgatcgttgg cgtcggcaaa 5340
cagtggtccg gcgcacgcgc tctggaggcc ttgctcacgg tggcgggaga gttgagaggt 5400
ccaccgttac agttggacac aggccaactt ctcaagattg caaaacgtgg cggcgtgacc 5460
gcagtggagg cagtgcatgc atggcgcaat gcactgacgg gtgccccgct caacttgacc 5520
ccccagcagg tggtggccat cgccagcaat aatggtggca agcaggcgct ggagacggtc 5580
cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccccagca ggtggtggcc 5640
atcgccagca atggcggtgg caagcaggcg ctggagacgg tccagcggct gttgccggtg 5700
ctgtgccagg cccacggctt gacccccccag caggtggtgg ccatcgccag caatggcggt 5760
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc 5820

```
ttgacccccc agcaggtggt ggccatcgcc agcaatggcg gtggcaagca ggcgctggag   5880
acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg   5940
gtggccatcg ccagcaatat tggtggcaag caggcgctgg agacggtgca ggcgctgttg   6000
ccggtgctgt gccaggccca cggcttgacc ccccagcagg tggtggccat cgccagcaat   6060
ggcggtggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc   6120
cacggcttga ccccccagca ggtggtggcc atcgccagca ataatggtgg caagcaggcg   6180
ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt gaccccccag   6240
caggtggtgg ccatcgccag caataatggt ggcaagcagg cgctggagac ggtccagcgg   6300
ctgttgccgg tgctgtgcca ggcccacggc ttgacccccc agcaggtggt ggccatcgcc   6360
agcaatggcg gtggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc   6420
caggcccacg gcttgacccc ccagcaggtg gtggccatcg ccagcaatgg cggtggcaag   6480
caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc   6540
ccggagcagg tggtggccat cgccagcaat attggtggca agcaggcgct ggagacggtg   6600
caggcgctgt tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc   6660
atcgccagcc acgatggcgg caagcaggcg ctggagacgg tccagcggct gttgccggtg   6720
ctgtgccagg cccacggctt gacccccccag caggtggtgg ccatcgccag caatggcggt   6780
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc   6840
ttgacccccc agcaggtggt ggccatcgcc agcaatggcg gtggcaagca ggcgctggag   6900
acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg   6960
gtggccatcg ccagcaatat tggtggcaag caggcgctgg agacggtgca ggcgctgttg   7020
ccggtgctgt gccaggccca cggcttgacc cctcagcagg tggtggccat cgccagcaat   7080
ggcggcggca ggccggcgct ggagagcatt gttgcccagt tatctcgccc tgatccgtcg   7140
gccgactgat aactcgagcg atcctctaga cgagctcctc gagcctgcag cagctgaagc   7200
tttggacttc ttcgccagag gtttggtcaa gtctccaatc aaggttgtcg gcttgtctac   7260
cttgccagaa atttacgaaa agatggaaaa gggtcaaatc gttggtagat acgttgttga   7320
cacttctaaa taagcgaatt tcttatgatt tatgattttt attattaaat aagttataaa   7380
aaaaataagt gtatacaaat tttaaagtga ctcttaggtt ttaaaacgaa aattcttatt   7440
cttgagtaac tctttcctgt aggtcaggtt gctttctcag gtatagcatg aggtcgctct   7500
tattgaccac acctctaccg gcatgcaagc ttggcgtaat catggtcata gctgtttcct   7560
gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag cataaagtgt   7620
aaagcctggg gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc   7680
gctttccagt cgggaaacct gtcgtgccag cagatctatt acattatggg tggtatgttg   7740
gaataaaaat caactatcat ctactaacta gtatttacgt tactagtata ttatcatata   7800
cggtgttaga agatgacgca aatgatgaga aatagtcatc taaattagtg gaagctgaaa   7860
cgcaaggatt gataatgtaa taggatcaat gaatattaac atataaaatg atgataataa   7920
tatttataga attgtgtaga attgcagatt ccctttatg gattcctaaa tcctcgagga   7980
gaacttctag tatatctaca tacctaatat tattgcctta ttaaaaatgg aatcccaaca   8040
attacatcaa aatccacatt ctcttcaaaa tcaattgtcc tgtacttcct tgttcatgtg   8100
tgttcaaaaa cgttatattt ataggataat tatactctat ttctcaacaa gtaattggtt   8160
gtttggccga gcggtctaag gcgcctgatt caagaaatat cttgaccgca gttaactgtg   8220
ggaatactca ggtatcgtaa gatgcaagag ttcgaatctc ttagcaacca ttatttttt   8280
cctcaacata acgagaacac acaggggcgc tatcgcacag aatcaaattc gatgactgga   8340
aatttttgt taatttcaga ggtcgcctga cgcatatacc tttttcaact gaaaaattgg   8400
gagaaaaagg aaaggtgaga gccgcggaac cggcttttca tatagaatag agaagcgttc   8460
atgactaaat gcttgcatca caatacttga agttgacaat attatttaag gacctattgt   8520
ttttccaat aggtggttag caatcgtctt actttctaac ttttcttacc ttttacattt   8580
cagcaatata tatatatata tttcaaggat ataccattct aatgtctgcc cctaagaaga   8640
tcgtcgtttt gccaggtgac cacgttggtc aagaaatcac agccgaagcc attaaggttc   8700
ttaaagctat ttctgatgtt cgttccaatg tcaagttcga tttcgaaaat catttaattg   8760
gtggtgctgc tatcgatgct acaggtgtcc cacttccaga tgaggcgctg gaagcctcca   8820
agaaggttga tgccgttttg ttaggtgctg tgggtggtcc taaatggggt accggtagtg   8880
ttagacctga acaaggttta ctaaaaatcc gtaaagaact tcaattgtac gccaacttaa   8940
gaccatgtaa ctttgcatcc gactctcttt tagacttatc tccaatcaag ccacaatttg   9000
ctaaaggtac tgacttcgtt gttgtcagag aattagtggg aggtatttac tttggtaaga   9060
gaaaggaaga cgatggtgat ggtgtcgctt gggatagtga acaatacacc gttccagaag   9120
tgcaaagaat cacaagaatg gccgctttca tggccctaca acatgagcca ccattgccta   9180
tttggtcctt ggataaagct aatgttttgg cctcttcaag attatggaga aaaactgtgg   9240
aggaaaccat caagaacgaa ttccctacat tgaaggttca acatcaattg attgattctg   9300
ccgccatgat cctagttaag aacccaaccc acctaaatgg tattataatc accagcaaca   9360
tgtttggtga tatcatctcc gatgaagcct ccgttatccc aggttccttg ggtttgttgc   9420
catctgcgtc cttggcctct ttgccagaca agaacaccgc atttggtttg tacgaaccat   9480
gccacggttc tgctccagat ttgccaaaga ataaggtcaa ccctatcgcc actatcttgt   9540
ctgctgcaat gatgttgaaa ttgtcattga acttgcctga agaaggtaag gccattgaag   9600
atgcagttaa aaaggttttg gatgcaggta tcagaactgg tgatttaggt ggttccaaca   9660
gtaccacgga agtcggtgat gctgtcgccg aagaagttaa gaaaatcctt gcttaaaaag   9720
attctctttt tttatgatat ttgtacataa actttataaa tgaaattcat aatagaaacg   9780
acacgaaatt acaaaatgga atatgttcat agggtagacg aaactatata cgcaatctac   9840
atacatttat caagaaggag aaaaaggagg atgtaaagga atacaggtaa gcaaattgat   9900
```

actaatggct caacgtgata aggaaaaaga attgcacttt aacattaata ttgacaagga 9960
ggagggcacc acacaaaaag ttaggtgtaa cagaaaatca tgaaactatg attcctaatt 10020
tatatattgg aggattttct ctaaaaaaaa aaaaatacaa caaataaaaa acactcaatg 10080
acctgaccat ttgatggagt ttaagtcaat accttcttga accatttccc ataatggtga 10140
aagttccctc aagaatttta ctctgtcaga aacggcctta acgacgtagt cgacctcctc 10200
ttcagtacta aatctaccaa taccaaatct gatggaagaa tgggctaatg catcatcctt 10260
acccagcgca tgtaaaacat aagaaggttc tagggaagca gatgtacagg ctgaacccga 10320
ggataatgcg atatccctta gtgccatcaa taaagattct ccttccacgt aggcgaaaga 10380
aacgttaaca caccctggat aacgatgatc tggagatccg ttcaacgtgg tatgttcagc 10440
ggataataga cctttgacta atttatcgga tagtcttttg atgtgagctt ggtcgttgtc 10500
aaattctttc ttcatcaatc tcgcagcttc accaaatccc gctaccaatg ggggggccaa 10560
agtaccagat ctgctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat 10620
tgggcgctct tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc ggctgcggcg 10680
agcggtatca gcatcgatga attccacgga ctatagacta tactagtata ctccgtctac 10740
tgtacgatac acttccgctc aggtccttgt cctttaacga ggccttacca ctcttttgtt 10800
actctattga tccagctcag caaaggcagt gtgatctaag attctatctt cgcgatgtag 10860
taaaactagc tagaccgaga aagagactag aaatgcaaaa ggcacttcta caatggctgc 10920
catcattatt atccgatgtg acgctgcagc ttctcaatga tattcgaata cgctttgagg 10980
agatacagcc taatatccga caaactgttt tacagattta cgatcgtact tgttacccat 11040
cattgaattt tgaacatccg aacctgggag ttttccctga aacagatagt atatttgaac 11100
ctgtataata atatatagtc tagcgcttta cggaagacaa tgtatgtatt tcggttcctg 11160
gagaaactat tgcatctatt gcataggtaa tcttgcacgt cgcatccccg gttcattttc 11220
tgcgtttcca tcttgcactt caatagcata tctttgttaa cgaagcatct gtgcttcatt 11280
ttgtagaaca aaaatgcaac gcgagagcgc taatttttca aacaaagaat ctgagctgca 11340
tttttacaga acagaaatgc aacgcgaaag cgctatttta ccaacgaaga atctgtgctt 11400
catttttgta aaacaaaaat gcaacgcgag agcgctaatt tttcaaacaa agaatctgag 11460
ctgcattttt acagaacaga aatgcaacgc gagagcgcta ttttaccaac aaagaatcta 11520
tacttctttt ttgttctaca aaaatgcatc ccgagagcgc tatttttcta acaaagcatc 11580
ttagattact ttttttctcc tttgtgcgct ctataatgca gtctcttgat aacttttttgc 11640
actgtaggtc cgttaaggtt agaagaaggc tactttggtg tctattttct cttccataaa 11700
aaaagcctga ctccacttcc cgcgtttact gattactagc gaagctgcgg gtgcattttt 11760
tcaagataaa ggcatccccg attatattct ataccgatgt ggattgcgca tactttgtga 11820
acagaaagtg atagcgttga tgattcttca ttggtcagaa aattatgaac ggtttcttct 11880
attttgtctc tatatactac gtataggaaa tgtttacatt ttcgtattgt tttcgattca 11940
ctctatgaat agttcttact acaatttttt tgtctaaaga gtaatactag agataaacat 12000
aaaaaatgta gaggtcgagt ttagatgcaa gttcaaggag cgaaaggtgg atgggtaggt 12060
tatataggga tatagcacag agatatatag caaagagata ctttgagca atgtttgtgg 12120
aagcggtatt cgcaatattt tagtagctcg ttacagtccg gtgcgttttt ggtttttga 12180
aagtgcgtct tcagagcgct tttggttttc aaaagcgctc tgaagttcct atactttcta 12240
gagaatagga acttcggaat aggaacttca aagcgtttcc gaaaacgagc gcttccgaaa 12300
atgcaacgcg agctgcgcac atacagctca ctgttcacgt cgcacctata tctgcgtgtt 12360
gcctgtatat atatatacat gagaagaacg gcatagtgcg tgtttatgct taaatgcgta 12420
cttatatgcg tctatttatg taggatgaaa ggtagtctag tacctcctgt gatattatcc 12480
cattccatgc ggggtatcgt atgcttcctt cagcactacc ctttagctgt tctatatgct 12540
gccactcctc aattggatta gtctcatcct tcaatgctat catttccttt gatattggat 12600
catatgcata gtaccgagaa actagtgcga agtagtgatc aggtattgct gttatctgat 12660
gagtatacgt tgtcctggcc acggcagaag cacgcttatc gctccaattt cccacaacat 12720
tagtcaactc cgttaggccc ttcattgaaa gaaatgaggt catcaaatgt cttccaatgt 12780
gagattttgg gccatttttt atagcaaaga ttgaataagg cgcatttttc ttcaaagctt 12840
tattgtacga tctgactaag ttatctttta ataattggta ttcctgttta ttgcttgaag 12900
aattgccggt cctatttact cgttttagga ctggttcaga attcatcgat gctcactcaa 12960
aggtcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa c 13011

<210> 363

<211> 13011
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS12110

<400> 363
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240

gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 720
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 780
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 840
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 900
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 960
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 1020
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 1080
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 1140
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 1200
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 1260
atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 1320
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 1380
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 1440
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 1500
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 1560
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 1620
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 1680
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 1740
atatttgaat gtatttagaa aaataaacaa atagggggttc cgcgcacatt tccccgaaaa 1800
gtgccacctg acgtccgatc aaaaatcatc gcttcgctga ttaattaccc cagaaataag 1860
gctaaaaaac taatcgcatt atcatcctat ggttgttaat ttgattcgtt catttgaagg 1920
tttgtggggc caggttactg ccaatttttc ctcttcataa ccataaaagc tagtattgta 1980
gaatctttat tgttcggagc agtgcggcgc gaggcacatc tgcgtttcag gaacgcgacc 2040
ggtgaagacg aggacgcacg gaggagagtc ttccttcgga gggctgtcac ccgctcggcg 2100
gcttctaatc cgtacttcaa tatagcaatg agcagttaag cgtattactg aaagttccaa 2160
agagaaggtt tttttaggct aatcgacctc gagcagatcc gccaggcgtg tatatagcgt 2220
ggatggccag gcaactttag tgctgacaca tacaggcata tatatatgtg tgcgacgaca 2280
catgatcata tggcatgcat gtgctctgta tgtatataaa actcttgttt tcttcttttc 2340
tctaaatatt ctttccttat acattaggtc ctttgtagca taaattacta tacttctata 2400
gacacgcaaa cacaaataca cagcggcctt gccaccatgg tggatctacg cacgctcggc 2460
tacagccagc agcaacagga gaagatcaaa ccgaaggttc gttcgacagt ggcgcagcac 2520
cacgaggcac tggtcggcca cgggtttaca cacgcgcaca tcgttgcgtt aagccaacac 2580
ccggcagcgt tagggaccgt cgctgtcaag tatcaggaca tgatcgcagc gttgccagag 2640
gcgacacacg aagcgatcgt tggcgtcggc aaacagtggt ccggcgcacg cgctctggag 2700
gccttgctca cggtggcggg agagttgaga ggtccaccgt tacagttgga cacaggccaa 2760
cttctcaaga ttgcaaaacg tggcggcgtg accgcagtgg aggcagtgca tgcatggcgc 2820
aatgcactga cgggtgcccc gctcaacttg acccccccagc aggtggtggc catcgccagc 2880
aataatggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 2940
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag 3000
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc 3060
cagcaggtgg tggccatcgc cagcaatggc ggtggcaagc aggcgctgga gacggtccag 3120
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 3180
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 3240
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 3300
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 3360
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 3420
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg 3480
gccatcgcca gcaataatgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg 3540
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaataat 3600
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 3660
ggcttgaccc cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg 3720
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 3780
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 3840
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc 3900
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag 3960
gcccacggct tgaccccgga gcaggtggtg gccatcgcca gccacgatgg cggcaagcag 4020
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc 4080
cagcaggtgg tggccatcgc cagcaatggc ggtggcaagc aggcgctgga gacggtccag 4140
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 4200
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 4260
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 4320

aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 4380
acccctcagc aggtggtggc catcgccagc aatggcggcg gcaggccggc gctggagagc 4440
attgttgccc agttatctcg ccctgatccg ggatcctccg gatctaaacg caacaaacca 4500
ggtaaggcca ctggcaaagg taaaccagtg aacaataagt ggctgaacaa cgctggcaaa 4560
gacttgggtt ctccagtgcc ggatcgcatt gccaacaagc tgagagacaa agaattcaag 4620
agttttgatg actttcgcgc aaaattctgg gaagaagtgt ctaaagatcc agaactgtct 4680
aaacagttta gtcgcaacaa caatgacaga atgaaagttg gtaaagctcc gaagacccgc 4740
acccaagacg tgtctggcaa acgcaccagt tttgaattgc atcatgaaaa accaatctct 4800
cagaatggtg gcgtttacga catggacaac atctctgtgg ttactccgaa aagacatatc 4860
gatattcatc gcggtaaagg atccaccggt ggatcctccg gagtggatct acgcacgctc 4920
ggctacagcc agcagcaaca ggagaagatc aaaccgaagg ttcgttcgac agtggcgcag 4980
caccacgagg cactggtcgg ccacgggttt acacacgcgc acatcgttgc gttaagccaa 5040
cacccggcag cgttagggac cgtcgctgtc aagtatcagg acatgatcgc agcgttgcca 5100
gaggcgacac acgaagcgat cgttggcgtc ggcaaacagt ggtccggcgc acgcgctctg 5160
gaggccttgc tcacggtggc gggagagttg agaggtccac cgttacagtt ggacacaggc 5220
caacttctca agattgcaaa acgtggcggc gtgaccgcag tggaggcagt gcatgcatgg 5280
cgcaatgcac tgacgggtgc cccgctcaac ttgaccccgg agcaggtggt ggccatcgcc 5340
agccacgatg gcggcaagca ggcgctggag acggtccagc ggctgttgcc ggtgctgtgc 5400
caggcccacg gcttgacccc ccagcaggtg gtggccatcg ccagcaatgg cggtggcaag 5460
caggcgctgg agacggtcca gcggctgttg ccggtgctgt gccaggccca cggcttgacc 5520
ccggagcagg tggtggccat cgccagcaat attggtggca agcaggcgct ggagacggtg 5580
caggcgctgt tgccggtgct gtgccaggcc cacggcttga ccccccagca ggtggtggcc 5640
atcgccagca atggcggtgg caagcaggcg ctggagacgg tccagcggct gttgccggtg 5700
ctgtgccagg cccacggctt gaccccggag caggtggtgg ccatcgccag caatattggt 5760
ggcaagcagg cgctggagac ggtgcaggcg ctgttgccgg tgctgtgcca ggcccacggc 5820
ttgaccccgg agcaggtggt ggccatcgcc agcaatattg gtggcaagca ggcgctggag 5880
acggtgcagg cgctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg 5940
gtggccatcg ccagcaatat tggtggcaag caggcgctgg agacggtgca ggcgctgttg 6000
ccggtgctgt gccaggccca cggcttgacc ccggagcagg tggtggccat cgccagccac 6060
gatggcggca agcaggcgct ggagacggtc cagcggctgt tgccggtgct gtgccaggcc 6120
cacggcttga ccccggagca ggtggtggcc atcgccagcc acgatggcgg caagcaggcg 6180
ctggagacgg tccagcggct gttgccggtg ctgtgccagg cccacggctt gacccccag 6240
caggtggtgg ccatcgccag caatggcggt ggcaagcagg cgctggagac ggtccagcgg 6300
ctgttgccgg tgctgtgcca ggcccacggc ttgaccccgg agcaggtggt ggccatcgcc 6360
agcaatattg gtggcaagca ggcgctggag acggtgcagg cgctgttgcc ggtgctgtgc 6420
caggcccacg gcttgacccc ggagcaggtg gtggccatcg ccagcaatat tggtggcaag 6480
caggcgctgg agacggtgca ggcgctgttg ccggtgctgt gccaggccca cggcttgacc 6540
ccggagcagg tggtggccat cgccagccac gatggcggca agcaggcgct ggagacggtc 6600
cagcggctgt tgccggtgct gtgccaggcc cacggcttga ccccggagca ggtggtggcc 6660
atcgccagcc acgatggcgg caagcaggcg ctggagacgg tccagcggct gttgccggtg 6720
ctgtgccagg cccacggctt gaccccggag caggtggtgg ccatcgccag ccacgatggc 6780
ggcaagcagg cgctggagac ggtccagcgg ctgttgccgg tgctgtgcca ggcccacggc 6840
ttgacccccc agcaggtggt ggccatcgcc agcaatggcg gtggcaagca ggcgctggag 6900
acggtccagc ggctgttgcc ggtgctgtgc caggcccacg gcttgacccc ggagcaggtg 6960
gtggccatcg ccagccacga tggcggcaag caggcgctgg agacggtcca gcggctgttg 7020
ccggtgctgt gccaggccca cggcttgacc cctcagcagg tggtggccat cgccagcaat 7080
ggcggcggca ggccggcgct ggagagcatt gttgcccagt tatctcgccc tgatccgtcg 7140
gccgactgat aactcgagcg atcctctaga cgagctcctc gagcctgcag cagctgaagc 7200
tttggacttc ttcgccagag gtttggtcaa gtctccaatc aaggttgtcg gcttgtctac 7260
cttgccagaa atttacgaaa agatggaaaa gggtcaaatc gttggtagat acgttgttga 7320
cacttctaaa taagcgaatt tcttatgatt tatgattttt attattaaat aagttataaa 7380
aaaaataagt gtatacaaat tttaaagtga ctcttaggtt ttaaaacgaa aattcttatt 7440
cttgagtaac tctttcctgt aggtcaggtt gctttctcag gtatagcatg aggtcgctct 7500
tattgaccac acctctaccg gcatgcaagc ttggcgtaat catggtcata gctgtttcct 7560
gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag cataaagtgt 7620
aaagcctggg gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc 7680
gctttccagt cgggaaacct gtcgtgccag cagatctatt acattatggg tggtatgttg 7740
gaataaaaat caactatcat ctactaacta gtatttacgt tactagtata ttatcatata 7800
cggtgttaga agatgacgca aatgatgaga aatagtcatc taaattagtg gaagctgaaa 7860
cgcaaggatt gataatgtaa taggatcaat gaatattaac atataaaatg atgataataa 7920
tatttataga attgtgtaga attgcagatt ccctttttatg gattcctaaa tcctcgagga 7980
gaacttctag tatatctaca tacctaatat tattgcctta ttaaaaatgg aatcccaaca 8040
attacatcaa aatccacatt ctcttcaaaa tcaattgtcc tgtacttcct tgttcatgtg 8100
tgttcaaaaa cgttatattt ataggataat tatactctat ttctcaacaa gtaattggtt 8160
gtttggccga gcggtctaag gcgcctgatt caagaaatat cttgaccgca gttaactgtg 8220
ggaatactca ggtatcgtaa gatgcaagag ttcgaatctc ttagcaacca ttatttttt 8280
cctcaacata acgagaacac acaggggcgc tatcgcacag aatcaaattc gatgactgga 8340
aatttttgt taatttcaga ggtcgcctga cgcatatacc tttttcaact gaaaaattgg 8400

```
gagaaaaagg aaaggtgaga gccgcggaac cggcttttca tatagaatag agaagcgttc   8460
atgactaaat gcttgcatca caatacttga agttgacaat attatttaag gacctattgt   8520
tttttccaat aggtggttag caatcgtctt actttctaac ttttcttacc ttttacattt   8580
cagcaatata tatatatata tttcaaggat ataccattct aatgtctgcc cctaagaaga   8640
tcgtcgtttt gccaggtgac cacgttggtc aagaaatcac agccgaagcc attaaggttc   8700
ttaaagctat ttctgatgtt cgttccaatg tcaagttcga tttcgaaaat catttaattg   8760
gtggtgctgc tatcgatgct acaggtgtcc cacttccaga tgaggcgctg gaagcctcca   8820
agaaggttga tgccgttttg ttaggtgctg tgggtggtcc taaatggggt accggtagtg   8880
ttagacctga acaaggttta ctaaaaatcc gtaaagaact tcaattgtac gccaacttaa   8940
gaccatgtaa ctttgcatcc gactctcttt tagacttatc tccaatcaag ccacaatttg   9000
ctaaaggtac tgacttcgtt gttgtcagag aattagtggg aggtatttac tttggtaaga   9060
gaaaggaaga cgatggtgat ggtgtcgctt gggatagtga acaatacacc gttccagaag   9120
tgcaaagaat cacaagaatg gccgctttca tggccctaca acatgagcca ccattgccta   9180
tttggtcctt ggataaagct aatgttttgg cctcttcaag attatggaga aaaactgtgg   9240
aggaaaccat caagaacgaa ttccctacat tgaaggttca acatcaattg attgattctg   9300
ccgccatgat cctagttaag aacccaaccc acctaaatgg tattataatc accagcaaca   9360
tgtttggtga tatcatctcc gatgaagcct ccgttatccc aggttccttg ggtttgttgc   9420
catctgcgtc cttggcctct ttgccagaca agaacaccgc atttggtttg tacgaaccat   9480
gccacggttc tgctccagat ttgccaaaga ataaggtcaa ccctatcgcc actatcttgt   9540
ctgctgcaat gatgttgaaa ttgtcattga acttgcctga agaaggtaag gccattgaag   9600
atgcagttaa aaaggttttg gatgcaggta tcagaactgg tgatttaggt ggttccaaca   9660
gtaccacgga agtcggtgat gctgtcgccg aagaagttaa gaaaatcctt gcttaaaaag   9720
attctctttt tttatgatat ttgtacataa actttataaa tgaaattcat aatagaaacg   9780
acacgaaatt acaaaatgga atatgttcat agggtagacg aaactatata cgcaatctac   9840
atacatttat caagaaggag aaaaaggagg atgtaaagga atacaggtaa gcaaattgat   9900
actaatggct caacgtgata aggaaaaaga attgcacttt aacattaata ttgacaagga   9960
ggagggcacc acacaaaaag ttaggtgtaa cagaaaatca tgaaactatg attcctaatt  10020
tatatattgg aggattttct ctaaaaaaaa aaaaatacaa caaataaaaa acactcaatg  10080
acctgaccat ttgatggagt ttaagtcaat accttcttga accatttccc ataatggtga  10140
aagttccctc aagaatttta ctctgtcaga aacggcctta acgacgtagt cgacctcctc  10200
ttcagtacta aatctaccaa taccaaatct gatggaagaa tgggctaatg catcatcctt  10260
acccagcgca tgtaaaacat aagaaggttc tagggaagca gatgtacagg ctgaacccga  10320
ggataatgcg atatccctta gtgccatcaa taaagattct ccttccacgt aggcgaaaga  10380
aacgttaaca caccctggat aacgatgatc tggagatccg ttcaacgtgg tatgttcagc  10440
ggataataga cctttgacta atttatcgga tagtcttttg atgtgagctt ggtcgttgtc  10500
aaattctttc ttcatcaatc tcgcagcttc accaaatccc gctaccaatg gggggccaa   10560
agtaccagat ctgctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat  10620
tgggcgctct tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc ggctgcggcg  10680
agcggtatca gcatcgatga attccacgga ctatagacta tactagtata ctccgtctac  10740
tgtacgatac acttccgctc aggtccttgt cctttaacga ggccttacca ctcttttgtt  10800
actctattga tccagctcag caaaggcagt gtgatctaag attctatctt cgcgatgtag  10860
taaaactagc tagaccgaga aagagactag aaatgcaaaa ggcacttcta caatggctgc  10920
catcattatt atccgatgtg acgctgcagc ttctcaatga tattcgaata cgctttgagg  10980
agatacagcc taatatccga caaactgttt tacagattta cgatcgtact tgttacccat  11040
cattgaattt tgaacatccg aacctgggag ttttccctga aacagatagt atatttgaac  11100
ctgtataata atatatagtc tagcgcttta cggaagacaa tgtatgtatt tcggttcctg  11160
gagaaactat tgcatctatt gcataggtaa tcttgcacgt cgcatccccg gttcattttc  11220
tgcgtttcca tcttgcactt caatagcata tctttgttaa cgaagcatct gtgcttcatt  11280
ttgtagaaca aaaatgcaac gcgagagcgc taatttttca aacaaagaat ctgagctgca  11340
tttttacaga acagaaatgc aacgcgaaag cgctatttta ccaacgaaga atctgtgctt  11400
cattttgta aaacaaaaat gcaacgcgag agcgctaatt tttcaaacaa agaatctgag  11460
ctgcattttt acagaacaga aatgcaacgc gagagcgcta ttttaccaac aaagaatcta  11520
tacttctttt ttgttctaca aaaatgcatc ccgagagcgc tatttttcta acaaagcatc  11580
ttagattact ttttttctcc tttgtgcgct ctataatgca gtctcttgat aacttttgc   11640
actgtaggtc cgttaaggtt agaagaaggc tactttggtg tctattttct cttccataaa  11700
aaaagcctga ctccacttcc cgcgtttact gattactagc gaagctgcgg gtgcattttt  11760
tcaagataaa ggcatccccg attatattct ataccgatgt ggattgcgca tactttgtga  11820
acagaaagtg atagcgttga tgattcttca ttggtcagaa aattatgaac ggtttcttct  11880
attttgtctc tatatactac gtataggaaa tgtttacatt ttcgtattgt tttcgattca  11940
ctctatgaat agttcttact acaatttttt tgtctaaaga gtaatactag agataaacat  12000
aaaaaatgta gaggtcgagt ttagatgcaa gttcaaggag cgaaaggtgg atgggtaggt  12060
tatataggga tatagcacag agatatatag caaagagata cttttgagca atgtttgtgg  12120
aagcggtatt cgcaatattt tagtagctcg ttacagtccg gtgcgttttt ggtttttga   12180
aagtgcgtct tcagagcgct tttggttttc aaaagcgctc tgaagttcct atactttcta  12240
gagaatagga acttcggaat aggaacttca aagcgtttcc gaaaacgagc gcttccgaaa  12300
atgcaacgcg agctgcgcac atacagctca ctgttcacgt cgcacctata tctgcgtgtt  12360
gcctgtatat atatatacat gagaagaacg gcatagtgcg tgtttatgct taaatgcgta  12420
cttatatgcg tctatttatg taggatgaaa ggtagtctag tacctcctgt gatattatcc  12480
```

cattccatgc ggggtatcgt atgcttcctt cagcactacc ctttagctgt tctatatgct 12540
gccactcctc aattggatta gtctcatcct tcaatgctat catttccttt gatattggat 12600
catatgcata gtaccgagaa actagtgcga agtagtgatc aggtattgct gttatctgat 12660
gagtatacgt tgtcctggcc acggcagaag cacgcttatc gctccaattt cccacaacat 12720
tagtcaactc cgttaggccc ttcattgaaa gaaatgaggt catcaaatgt cttccaatgt 12780
gagattttgg gccatttttt atagcaaaga ttgaataagg cgcatttttc ttcaaagctt 12840
tattgtacga tctgactaag ttatctttta ataattggta ttcctgttta ttgcttgaag 12900
aattgccggt cctatttact cgttttagga ctggttcaga attcatcgat gctcactcaa 12960
aggtcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa c 13011

<210> 364

<211> 5594
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS15802

<400> 364
ggtctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc 60
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 120
gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 180
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 240
ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 300
attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcgc 360
gtcaattaac cctcactaaa gggaacaaaa gctgttaatt aagcgcgcca cacgcaaaca 420
caaatacaca gcggccttgc caccatggtg gatctacgca cgctcggcta cagccagcag 480
caacaggaga agatcaaacc gaaggttcgt tcgacagtgg cgcagcacca cgaggcactg 540
gtcggccacg ggtttacaca cgcgcacatc gttgcgttaa gccaacaccc ggcagcgtta 600
gggaccgtcg ctgtcaagta tcaggacatg atcgcagcgt tgccagaggc gacacacgaa 660
gcgatcgttg gcgtcggcaa acagtggtcc ggcgcacgcg ctctggaggc cttgctcacg 720
gtggcgggag agttgagagg tccaccgtta cagttggaca caggccaact tctcaagatt 780
gcaaaacgtg gcggcgtgac cgcagtggag gcagtgcatg catggcgcaa tgcactgacg 840
ggtgccccgc tcaacttgac cccccagcag gtggtggcca tcgccagcaa taatggtggc 900
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg 960
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg 1020
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg 1080
gccatcgcca gcaatggcgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg 1140
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 1200
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 1260
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg 1320
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag 1380
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg 1440
ttgccggtgc tgtgccaggc ccacggcttg acccccccagc aggtggtggc catcgccagc 1500
aataatggtg gcaagcaggc gctggagacg gtccagcggc tgttgccggt gctgtgccag 1560
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaataatgg tggcaagcag 1620
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc 1680
cagcaggtgg tggccatcgc cagcaatggc ggtggcaagc aggcgctgga gacggtccag 1740
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc 1800
gccagcaatg gcggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg 1860
tgccaggccc acggcttgac cccggagcag gtggtggcca tcgccagcaa tattggtggc 1920
aagcaggcgc tggagacggt gcaggcgctg ttgccggtgc tgtgccaggc ccacggcttg 1980
accccggagc aggtggtggc catcgccagc cacgatggcg gcaagcaggc gctggagacg 2040
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgacccccca gcaggtggtg 2100
gccatcgcca gcaatggcgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg 2160
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc 2220
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac 2280
ggcttgaccc cggagcaggt ggtggccatc gccagcaata ttggtggcaa gcaggcgctg 2340
gagacggtgc aggcgctgtt gccggtgctg tgccaggccc acggcttgac ccctcagcag 2400
gtggtggcca tcgccagcaa tggcggcggc aggccggcgc tggagagcat tgttgcccag 2460
ttatctcgcc ctgatccggg atccaagtct ggcatctacc agattaagaa caccctgaat 2520
aacaaagtgt atgttggatc tgccaaggac tttgagaaaa gatggaagag gcacttcaag 2580
gatctggaaa agggttgcca cagctccatc aagctccaaa ggtctttcaa caaacatggc 2640
aatgtctttg agtgctccat tttggaagag atcccttatg aaaaggacct cattatagag 2700
agggaaaact tctggatcaa ggagctgaac tcaaaaatca atggatacaa cattgcagat 2760
gccacatttg gtgacacctg ttctactcac ccccttaagg aagagatcat taaaaagagg 2820

tctgagactg tgaaggctaa gatgttgaaa ctggggcctg atggcagaaa ggcactctac 2880
tctaagccag gtagcaaaaa tggaaggtgg aaccctgaga cccacaagtt ctgcaaatgt 2940
ggggtgagaa tacagacatc agcctacact tgcagcaagt gtaggaacag atcaggagag 3000
aataactcct ttttcaacca caagcattct caaggtccat cggccgactg ataagtcgac 3060
caccaccacc accaccactg ataagagctc ctcgagcctg cagcagctga agctttggac 3120
ttctggcgcg cccaattcgc cctatagtga gtcgtattac gtcgcgctca ctggccgtcg 3180
ttttacaacg tcgtgactgg gaaaaccctg gcgttaccca acttaatcgc cttgcagcac 3240
atccccctttt cgccagctgg cgtaatagcg aagaggcccg caccgaaacg cccttcccaa 3300
cagttgcgca gcctgaatgg cgaatgggag cgccctgtag cggcgcatta agcgcggcgg 3360
gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt 3420
tcgctttctt cccttccttt ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc 3480
gggggctccc tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg 3540
attagggtga tggttggcct gtagtgggcc atagccctga tagacggttt ttcgcccttt 3600
gacgttggag tccacgttct ttaatagtgg actcttgttc caaactggaa caacactcaa 3660
ccctatctcg gtctattctt ttgatttata agggattttg ccgatttcgg cctattggtt 3720
aaaaaatgag ctgatttaac aaaaatttaa cgcgaatttt aacaaaatat taacgcttac 3780
aatttaggtg gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt atttttctaa 3840
atacattcaa atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat 3900
tgaaaaagga agagtatgag tattcaacat ttccgtgtcg cccttattcc ctttttgcg 3960
gcattttgcc ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa 4020
gatcagttgg gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt 4080
gagagttttc gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt 4140
ggcgcggtat tatcccgtat tgacgccggg caagagcaac tcggtcgccg catacactat 4200
tctcagaatg acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg 4260
acagtaagag aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta 4320
cttctgacaa cgatcggagg accgaaggag ctaaccgctt ttttgcacaa catgggggat 4380
catgtaactc gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag 4440
cgtgacacca cgatgcctgt agcaatggca acaacgttgc gcaaactatt aactggcgaa 4500
ctacttactc tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca 4560
ggaccacttc tgcgctcggc ccttccggct ggctggttta ttgctgataa atctggagcc 4620
ggtgagcgtg gttctcgcgg tatcattgca gcactggggc cagatggtaa gccctcccgt 4680
atcgtagtta tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc 4740
gctgagatag gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat 4800
atactttaga ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt 4860
tttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac 4920
cccgtagaaa agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc 4980
ttgcaaacaa aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca 5040
actcttttc cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta 5100
gtgtagccgt agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct 5160
ctgctaatcc tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg 5220
gactcaagac gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc 5280
acacagccca gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta 5340
tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg 5400
gtcggaacag gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt 5460
cctgtcgggt ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcaggggggg 5520
cggagcctat ggaaaaacgc cagcaacgcg gccttttac ggttcctggc cttttgctgg 5580
ccttttgctc acat 5594

<210> 365

<211> 8588
<212> DNA
<213> Artificial sequence

<220>
<223> pCLS8990

<400> 365
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 60
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 120
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac 360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 420
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 480
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 540

```
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt    600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg    660
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc    720
atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca    780
tttccccgaa aagtgccacc tgacgtccga tcaaaaatca tcgcttcgct gattaattac    840
cccagaaata aggctaaaaa actaatcgca ttatcatcct atggttgtta atttgattcg    900
ttcatttgaa ggtttgtggg gccaggttac tgccaatttt tcctcttcat aaccataaaa    960
gctagtattg tagaatcttt attgttcgga gcagtgcggc gcgaggcaca tctgcgtttc   1020
aggaacgcga ccggtgaaga cgaggacgca cggaggagag tcttccttcg gagggctgtc   1080
acccgctcgg cggcttctaa tccgtacttc aatatagcaa tgagcagtta agcgtattac   1140
tgaaagttcc aaagagaagg ttttttttagg ctaatcgacc tcgagcagat ccgccaggcg   1200
tgtatatagc gtggatggcc aggcaacttt agtgctgaca catacaggca tatatatatg   1260
tgtgcgacga cacatgatca tatggcatgc atgtgctctg tatgtatata aaactcttgt   1320
tttcttcttt tctctaaata ttctttcctt atacattagg tcctttgtag cataaattac   1380
tatacttcta tagacacgca aacacaaata cacagcggcc ttgccaccat ggtggatcta   1440
cgcacgctcg gctacagcca gcagcaacag gagaagatca aaccgaaggt tcgttcgaca   1500
gtggcgcagc accacgaggc actggtcggc cacgggttta cacacgcgca catcgttgcg   1560
ttaagccaac acccggcagc gttagggacc gtcgctgtca agtatcagga catgatcgca   1620
gcgttgccag aggcgacaca cgaagcgatc gttggcgtcg gcaaacagtg gtccggcgca   1680
cgcgctctgg aggccttgct cacggtggcg ggagagttga gaggtccacc gttacagttg   1740
gacacaggcc aacttctcaa gattgcaaaa cgtggcggcg tgaccgcagt ggaggcagtg   1800
catgcatggc gcaatgcact gacgggtgcc ccgctcaact tgacccccca gcaggtggtg   1860
gccatcgcca gcaataatgg tggcaagcag gcgctggaga cggtccagcg gctgttgccg   1920
gtgctgtgcc aggcccacgg cttgaccccc cagcaggtgg tggccatcgc cagcaatggc   1980
ggtggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac   2040
ggcttgaccc cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg   2100
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag   2160
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg   2220
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc   2280
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag   2340
gcccacggct tgacccccca gcaggtggtg gccatcgcca gcaatggcgg tggcaagcag   2400
gcgctggaga cggtccagcg gctgttgccg gtgctgtgcc aggcccacgg cttgaccccc   2460
cagcaggtgg tggccatcgc cagcaataat ggtggcaagc aggcgctgga gacggtccag   2520
cggctgttgc cggtgctgtg ccaggcccac ggcttgaccc cccagcaggt ggtggccatc   2580
gccagcaata atggtggcaa gcaggcgctg gagacggtcc agcggctgtt gccggtgctg   2640
tgccaggccc acggcttgac cccccagcag gtggtggcca tcgccagcaa tggcggtggc   2700
aagcaggcgc tggagacggt ccagcggctg ttgccggtgc tgtgccaggc ccacggcttg   2760
acccccccagc aggtggtggc catcgccagc aatggcggtg gcaagcaggc gctggagacg   2820
gtccagcggc tgttgccggt gctgtgccag gcccacggct tgaccccgga gcaggtggtg   2880
gccatcgcca gcaatattgg tggcaagcag gcgctggaga cggtgcaggc gctgttgccg   2940
gtgctgtgcc aggcccacgg cttgaccccg gagcaggtgg tggccatcgc cagccacgat   3000
ggcggcaagc aggcgctgga gacggtccag cggctgttgc cggtgctgtg ccaggcccac   3060
ggcttgaccc cccagcaggt ggtggccatc gccagcaatg gcggtggcaa gcaggcgctg   3120
gagacggtcc agcggctgtt gccggtgctg tgccaggccc acggcttgac cccccagcag   3180
gtggtggcca tcgccagcaa tggcggtggc aagcaggcgc tggagacggt ccagcggctg   3240
ttgccggtgc tgtgccaggc ccacggcttg accccggagc aggtggtggc catcgccagc   3300
aatattggtg gcaagcaggc gctggagacg gtgcaggcgc tgttgccggt gctgtgccag   3360
gcccacggct tgacccctca gcaggtggtg gccatcgcca gcaatggcgg cggcaggccg   3420
gcgctggaga gcattgttgc ccagttatct cgccctgatc cgggatccaa gtctggcatc   3480
taccagatta agaacaccct gaataacaaa gtgtatgttg gatctgccaa ggactttgag   3540
aaaagatgga agaggcactt caaggatctg gaaaagggtt gccacagctc catcaagctc   3600
caaaggtctt tcaacaaaca tggcaatgtc tttgagtgct ccattttgga agagatccct   3660
tatgaaaagg acctcattat agagagggaa aacttctgga tcaaggagct gaactcaaaa   3720
atcaatggat acaacattgc agatgccaca tttggtgaca cctgttctac tcacccccctt   3780
aaggaagaga tcattaaaaa gaggtctgag actgtgaagg ctaagatgtt gaaactgggg   3840
cctgatggca gaaaggcact ctactctaag ccaggtagca aaaatggaag gtggaaccct   3900
gagacccaca agttctgcaa atgtggggtg agaatacaga catcagccta cacttgcagc   3960
aagtgtagga acagatcagg agagaataac tccttttca accacaagca ttctcaaggt   4020
ccatcggccg actgataact cgagcgatcc tctagacgag ctcctcgagc ctgcagcagc   4080
tgaagctttg gacttcttcg ccagaggttt ggtcaagtct ccaatcaagg ttgtcggctt   4140
gtctaccttg ccagaaattt acgaaaagat ggaaaagggt caaatcgttg gtagatacgt   4200
tgttgacact tctaaataag cgaatttctt atgatttatg atttttatta ttaaataagt   4260
tataaaaaaa ataagtgtat acaaatttta aagtgactct taggttttaa aacgaaaatt   4320
cttattcttg agtaactctt tcctgtaggt caggttgctt tctcaggtat agcatgaggt   4380
cgctcttatt gaccacacct ctaccggcat gcaagcttgg cgtaatcatg gtcatagctg   4440
tttcctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata   4500
aagtgtaaag cctggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca   4560
ctgcccgctt tccagtcggg aaacctgtcg tgccagcaga tctgtttagc ttgcctcgtc   4620
```

cccgccgggt cacccggcca gcgacatgga ggcccagaat accctccttg acagtcttga 4680
cgtgcgcagc tcaggggcat gatgtgactg tcgcccgtac atttagccca tacatcccca 4740
tgtataatca tttgcatcca tacattttga tggccgcacg gcgcgaagca aaaattacgg 4800
ctcctcgctg cagacctgcg agcagggaaa cgctcccctc acagacgcgt tgaattgtcc 4860
ccacgccgcg cccctgtaga gaaatataaa aggttaggat ttgccactga ggttcttctt 4920
tcatatactt ccttttaaaa tcttgctagg atacagttct cacatcacat ccgaacataa 4980
acaaccatgc atgggtaagg aaaagactca cgtttcgagg ccgcgattaa attccaacat 5040
ggatgctgat ttatatgggt ataaatgggc tcgcgataat gtcgggcaat caggtgcgac 5100
aatctatcga ttgtatggga agcccgatgc gccagagttg tttctgaaac atggcaaagg 5160
tagcgttgcc aatgatgtta cagatgagat ggtcagacta aactggctga cggaatttat 5220
gcctcttccg accatcaagc attttatccg tactcctgat gatgcatggt tactcaccac 5280
tgcgatcccc ggcaaaacag cattccaggt attagaagaa tatcctgatt caggtgaaaa 5340
tattgttgat gcgctggcag tgttcctgcg ccggttgcat tcgattcctg tttgtaattg 5400
tccttttaac agcgatcgcg tatttcgcct cgctcaggcg caatcacgaa tgaataacgg 5460
tttggttgat gcgagtgatt ttgatgacga gcgtaatggc tggcctgttg aacaagtctg 5520
gaaagaaatg cataagcttt tgccattctc accggattca gtcgtcactc atggtgattt 5580
ctcacttgat aaccttattt ttgacgaggg gaaattaata ggttgtattg atgttggacg 5640
agtcggaatc gcagaccgat accaggatct tgccatccta tggaactgcc tcggtgagtt 5700
ttctccttca ttacagaaac ggctttttca aaaatatggt attgataatc ctgatatgaa 5760
taaattgcag tttcatttga tgctcgatga gtttttctaa tcagtactga caataaaaag 5820
attcttgttt tcaagaactt gtcatttgta tagttttttt atattgtagt tgttctattt 5880
taatcaaatg ttagcgtgat ttatattttt tttcgcctcg acatcatctg cccagatgcg 5940
aagttaagtg cgcagaaagt aatatcatgc gtcaatcgta tgtgaatgct ggtcgctata 6000
ctgctgtcga ttcgatacta acgccgccat ccagtgtcga aaacgagctc gaattcatcg 6060
atgatatcag atccactagt ggcctatgcg accgcggatc tgccggtctc cctatagtga 6120
gtcgtattaa tttcgataag ccaggttaac ctgcattaat gaatcggcca acgcgcgggg 6180
agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg 6240
gtcgttcggc tgcggcgagc ggtatcagca tcgatgaatt ccacggacta tagactatac 6300
tagtatactc cgtctactgt acgatacact tccgctcagg tccttgtcct ttaacgaggc 6360
cttaccactc ttttgttact ctattgatcc agctcagcaa aggcagtgtg atctaagatt 6420
ctatcttcgc gatgtagtaa aactagctag accgagaaag agactagaaa tgcaaaaggc 6480
acttctacaa tggctgccat cattattatc cgatgtgacg ctgcagcttc tcaatgatat 6540
tcgaatacgc tttgaggaga tacagcctaa tatccgacaa actgttttac agatttacga 6600
tcgtacttgt tacccatcat tgaattttga acatccgaac ctgggagttt tccctgaaac 6660
agatagtata tttgaacctg tataataata tatagtctag cgctttacgg aagacaatgt 6720
atgtatttcg gttcctggag aaactattgc atctattgca taggtaatct tgcacgtcgc 6780
atccccggtt cattttctgc gtttccatct tgcacttcaa tagcatatct ttgttaacga 6840
agcatctgtg cttcattttg tagaacaaaa atgcaacgcg agagcgctaa tttttcaaac 6900
aaagaatctg agctgcattt ttacagaaca gaaatgcaac gcgaaagcgc tattttacca 6960
acgaagaatc tgtgcttcat ttttgtaaaa caaaaatgca acgcgagagc gctaattttt 7020
caaacaaaga atctgagctg cattttaca gaacagaaat gcaacgcgag agcgctattt 7080
taccaacaaa gaatctatac ttctttttg ttctacaaaa atgcatcccg agagcgctat 7140
ttttctaaca aagcatctta gattactttt tttctccttt gtgcgctcta taatgcagtc 7200
tcttgataac tttttgcact gtaggtccgt taaggttaga agaaggctac tttggtgtct 7260
attttctctt ccataaaaaa agcctgactc cacttcccgc gtttactgat tactagcgaa 7320
gctgcgggtg catttttca agataaaggc atccccgatt atattctata ccgatgtgga 7380
ttgcgcatac tttgtgaaca gaaagtgata gcgttgatga ttcttcattg gtcagaaaat 7440
tatgaacggt ttcttctatt ttgtctctat atactacgta taggaaatgt ttacattttc 7500
gtattgtttt cgattcactc tatgaatagt tcttactaca attttttttgt ctaaagagta 7560
atactagaga taaacataaa aaatgtagag gtcgagttta gatgcaagtt caaggagcga 7620
aaggtggatg ggtaggttat atagggatat agcacagaga tatatagcaa agagatactt 7680
ttgagcaatg tttgtggaag cggtattcgc aatattttag tagctcgtta cagtccggtg 7740
cgtttttggt tttttgaaag tgcgtcttca gagcgctttt ggttttcaaa agcgctctga 7800
agttcctata ctttctagag aataggaact tcggaatagg aacttcaaag cgtttccgaa 7860
aacgagcgct tccgaaaatg caacgcgagc tgcgcacata cagctcactg ttcacgtcgc 7920
acctatatct gcgtgttgcc tgtatatata tatacatgag aagaacggca tagtgcgtgt 7980
ttatgcttaa atgcgtactt atatgcgtct atttatgtag gatgaaaggt agtctagtac 8040
ctcctgtgat attatcccat tccatgcggg gtatcgtatg cttccttcag cactaccctt 8100
tagctgttct atatgctgcc actcctcaat tggattagtc tcatccttca atgctatcat 8160
ttcctttgat attggatcat atgcatagta ccgagaaact agtgcgaagt agtgatcagg 8220
tattgctgtt atctgatgag tatacgttgt cctggccacg gcagaagcac gcttatcgct 8280
ccaatttccc acaacattag tcaactccgt taggcccttc attgaaagaa atgaggtcat 8340
caaatgtctt ccaatgtgag attttgggcc atttttata gcaaagattg aataaggcgc 8400
atttttcttc aaagctttat tgtacgatct gactaagtta tcttttaata attggtattc 8460
ctgtttattg cttgaagaat tgccggtcct atttactcgt tttaggactg gttcagaatt 8520
catcgatgct cactcaaagg tcggtaatac ggttatccac agaatcaggg gataacgcag 8580
gaaagaac 8588

<210> 366

<211> 582
<212> PRT
<213> Artificial sequence

<220>
<223> Colicin-E9

<400> 366

```
Met Ser Gly Gly Asp Gly Arg Gly His Asn Thr Gly Ala His Ser Thr
1               5                   10                  15

Ser Gly Asn Ile Asn Gly Gly Pro Thr Gly Ile Gly Val Ser Gly Gly
            20                  25                  30

Ala Ser Asp Gly Ser Gly Trp Ser Ser Glu Asn Asn Pro Trp Gly Gly
        35                  40                  45

Gly Ser Gly Ser Gly Ile His Trp Gly Gly Gly Ser Gly Arg Gly Asn
    50                  55                  60

Gly Gly Gly Asn Gly Asn Ser Gly Gly Gly Ser Gly Thr Gly Gly Asn
65                  70                  75                  80

Leu Ser Ala Val Ala Ala Pro Val Ala Phe Gly Phe Pro Ala Leu Ser
                85                  90                  95

Thr Pro Gly Ala Gly Gly Leu Ala Val Ser Ile Ser Ala Ser Glu Leu
            100                 105                 110

Ser Ala Ala Ile Ala Gly Ile Ile Ala Lys Leu Lys Lys Val Asn Leu
        115                 120                 125

Lys Phe Thr Pro Phe Gly Val Val Leu Ser Ser Leu Ile Pro Ser Glu
    130                 135                 140

Ile Ala Lys Asp Asp Pro Asn Met Met Ser Lys Ile Val Thr Ser Leu
145                 150                 155                 160

Pro Ala Asp Asp Ile Thr Glu Ser Pro Val Ser Ser Leu Pro Leu Asp
                165                 170                 175

Lys Ala Thr Val Asn Val Asn Val Arg Val Val Asp Asp Val Lys Asp
            180                 185                 190

Glu Arg Gln Asn Ile Ser Val Val Ser Gly Val Pro Met Ser Val Pro
        195                 200                 205

Val Val Asp Ala Lys Pro Thr Glu Arg Pro Gly Val Phe Thr Ala Ser
    210                 215                 220

Ile Pro Gly Ala Pro Val Leu Asn Ile Ser Val Asn Asp Ser Thr Pro
225                 230                 235                 240
```

```
Ala Val Gln Thr Leu Ser Pro Gly Val Thr Asn Asn Thr Asp Lys Asp
                245                 250                 255

Val Arg Pro Ala Gly Phe Thr Gln Gly Gly Asn Thr Arg Asp Ala Val
            260                 265                 270

Ile Arg Phe Pro Lys Asp Ser Gly His Asn Ala Val Tyr Val Ser Val
        275                 280                 285

Ser Asp Val Leu Ser Pro Asp Gln Val Lys Gln Arg Gln Asp Glu Glu
    290                 295                 300

Asn Arg Arg Gln Gln Glu Trp Asp Ala Thr His Pro Val Glu Ala Ala
305                 310                 315                 320

Glu Arg Asn Tyr Glu Arg Ala Arg Ala Glu Leu Asn Gln Ala Asn Glu
                325                 330                 335

Asp Val Ala Arg Asn Gln Glu Arg Gln Ala Lys Ala Val Gln Val Tyr
            340                 345                 350

Asn Ser Arg Lys Ser Glu Leu Asp Ala Ala Asn Lys Thr Leu Ala Asp
        355                 360                 365

Ala Ile Ala Glu Ile Lys Gln Phe Asn Arg Phe Ala His Asp Pro Met
    370                 375                 380

Ala Gly Gly His Arg Met Trp Gln Met Ala Gly Leu Lys Ala Gln Arg
385                 390                 395                 400

Ala Gln Thr Asp Val Asn Asn Lys Gln Ala Ala Phe Asp Ala Ala Ala
                405                 410                 415

Lys Glu Lys Ser Asp Ala Asp Ala Ala Leu Ser Ala Ala Gln Glu Arg
            420                 425                 430

Arg Lys Gln Lys Glu Asn Lys Glu Lys Asp Ala Lys Asp Lys Leu Asp
        435                 440                 445

Lys Glu Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys Gly Lys
    450                 455                 460

Pro Val Gly Asp Lys Trp Leu Asp Asp Ala Gly Lys Asp Ser Gly Ala
465                 470                 475                 480

Pro Ile Pro Asp Arg Ile Ala Asp Lys Leu Arg Asp Lys Glu Phe Lys
                485                 490                 495

Ser Phe Asp Asp Phe Arg Lys Ala Val Trp Glu Glu Val Ser Lys Asp
            500                 505                 510
```

```
Pro Glu Leu Ser Lys Asn Leu Asn Pro Ser Asn Lys Ser Ser Val Ser
        515                 520                 525

Lys Gly Tyr Ser Pro Phe Thr Pro Lys Asn Gln Gln Val Gly Gly Arg
    530                 535                 540

Lys Val Tyr Glu Leu His His Asp Lys Pro Ile Ser Gln Gly Gly Glu
545                 550                 555                 560

Val Tyr Asp Met Asp Asn Ile Arg Val Thr Thr Pro Lys Arg His Ile
                565                 570                 575

Asp Ile His Arg Gly Lys
            580
```

<210> 367

<211> 511
<212> PRT
<213> Artificial sequence

<220>
<223> APFL

<400> 367

```
Met Ser Gly Gly Phe Glu Leu Gln Pro Arg Asp Gly Gly Pro Arg Val
1               5                   10                  15

Ala Leu Ala Pro Gly Glu Thr Val Ile Gly Arg Gly Pro Leu Leu Gly
            20                  25                  30

Ile Thr Asp Lys Arg Val Ser Arg Arg His Ala Ile Leu Glu Val Ala
        35                  40                  45

Gly Gly Gln Leu Arg Ile Lys Pro Ile His Thr Asn Pro Cys Phe Tyr
    50                  55                  60

Gln Ser Ser Glu Lys Ser Gln Leu Leu Pro Leu Lys Pro Asn Leu Trp
65                  70                  75                  80

Cys Tyr Leu Asn Pro Gly Asp Ser Phe Ser Leu Leu Val Asp Lys Tyr
                85                  90                  95

Ile Phe Arg Ile Leu Ser Ile Pro Ser Glu Val Glu Met Gln Cys Thr
            100                 105                 110

Leu Arg Asn Ser Gln Val Leu Asp Glu Asp Asn Ile Leu Asn Glu Thr
        115                 120                 125

Pro Lys Ser Pro Val Ile Asn Leu Pro His Glu Thr Thr Gly Ala Ser
    130                 135                 140
```

```
Gln Leu Glu Gly Ser Thr Glu Ile Ala Lys Thr Gln Met Thr Pro Thr
145                 150                 155                 160

Asn Ser Val Ser Phe Leu Gly Glu Asn Arg Asp Cys Asn Lys Gln Gln
                165                 170                 175

Pro Ile Leu Ala Glu Arg Lys Arg Ile Leu Pro Thr Trp Met Leu Ala
            180                 185                 190

Glu His Leu Ser Asp Gln Asn Leu Ser Val Pro Ala Ile Ser Gly Gly
        195                 200                 205

Asn Val Ile Gln Gly Ser Gly Lys Glu Glu Ile Cys Lys Asp Lys Ser
    210                 215                 220

Gln Leu Asn Thr Thr Gln Gln Gly Arg Arg Gln Leu Ile Ser Ser Gly
225                 230                 235                 240

Ser Ser Glu Asn Thr Ser Ala Glu Gln Asp Thr Gly Glu Glu Cys Lys
                245                 250                 255

Asn Thr Asp Gln Glu Glu Ser Thr Ile Ser Ser Lys Glu Met Pro Gln
            260                 265                 270

Ser Phe Ser Ala Ile Thr Leu Ser Asn Thr Glu Met Asn Asn Ile Lys
        275                 280                 285

Thr Asn Ala Gln Arg Asn Lys Leu Pro Ile Glu Glu Leu Gly Lys Val
    290                 295                 300

Ser Lys His Lys Ile Ala Thr Lys Arg Thr Pro His Lys Glu Asp Glu
305                 310                 315                 320

Ala Met Ser Cys Ser Glu Asn Cys Ser Ser Ala Gln Gly Asp Ser Leu
                325                 330                 335

Gln Asp Glu Ser Gln Gly Ser His Ser Glu Ser Ser Ser Asn Pro Ser
            340                 345                 350

Asn Pro Glu Thr Leu His Ala Lys Ala Thr Asp Ser Val Leu Gln Gly
        355                 360                 365

Ser Glu Gly Asn Lys Val Lys Arg Thr Ser Cys Met Tyr Gly Ala Asn
    370                 375                 380

Cys Tyr Arg Lys Asn Pro Val His Phe Gln His Phe Ser His Pro Gly
385                 390                 395                 400

Asp Ser Asp Tyr Gly Gly Val Gln Ile Val Gly Gln Asp Glu Thr Asp
                405                 410                 415
```

Asp Arg Pro Glu Cys Pro Tyr Gly Pro Ser Cys Tyr Arg Lys Asn Pro
420 425 430

Gln His Lys Ile Glu Tyr Arg His Asn Thr Leu Pro Val Arg Asn Val
435 440 445

Leu Asp Glu Asp Asn Asp Asn Val Gly Gln Pro Asn Glu Tyr Asp Leu
450 455 460

Asn Asp Ser Phe Leu Asp Asp Glu Glu Glu Asp Tyr Glu Pro Thr Asp
465 470 475 480

Glu Asp Ser Asp Trp Glu Pro Gly Lys Glu Asp Glu Glu Lys Glu Asp
485 490 495

Val Glu Glu Leu Leu Lys Glu Ala Lys Arg Phe Met Lys Arg Lys
500 505 510

<210> 368

<211> 583
<212> PRT
<213> Artificial sequence

<220>
<223> FokI wild-type

<400> 368
Met Phe Leu Ser Met Val Ser Lys Ile Arg Thr Phe Gly Trp Val Gln
1 5 10 15

Asn Pro Gly Lys Phe Glu Asn Leu Lys Arg Val Val Gln Val Phe Asp
20 25 30

Arg Asn Ser Lys Val His Asn Glu Val Lys Asn Ile Lys Ile Pro Thr
35 40 45

Leu Val Lys Glu Ser Lys Ile Gln Lys Glu Leu Val Ala Ile Met Asn
50 55 60

Gln His Asp Leu Ile Tyr Thr Tyr Lys Glu Leu Val Gly Thr Gly Thr
65 70 75 80

Ser Ile Arg Ser Glu Ala Pro Cys Asp Ala Ile Ile Gln Ala Thr Ile
85 90 95

Ala Asp Gln Gly Asn Lys Lys Gly Tyr Ile Asp Asn Trp Ser Ser Asp
100 105 110

Gly Phe Leu Arg Trp Ala His Ala Leu Gly Phe Ile Glu Tyr Ile Asn
115 120 125

Lys Ser Asp Ser Phe Val Ile Thr Asp Val Gly Leu Ala Tyr Ser Lys
130 135 140

```
Ser Ala Asp Gly Ser Ala Ile Glu Lys Glu Ile Leu Ile Glu Ala Ile
145                 150                 155                 160

Ser Ser Tyr Pro Pro Ala Ile Arg Ile Leu Thr Leu Leu Glu Asp Gly
                165                 170                 175

Gln His Leu Thr Lys Phe Asp Leu Gly Lys Asn Leu Gly Phe Ser Gly
            180                 185                 190

Glu Ser Gly Phe Thr Ser Leu Pro Glu Gly Ile Leu Leu Asp Thr Leu
        195                 200                 205

Ala Asn Ala Met Pro Lys Asp Lys Gly Glu Ile Arg Asn Asn Trp Glu
    210                 215                 220

Gly Ser Ser Asp Lys Tyr Ala Arg Met Ile Gly Gly Trp Leu Asp Lys
225                 230                 235                 240

Leu Gly Leu Val Lys Gln Gly Lys Lys Glu Phe Ile Ile Pro Thr Leu
                245                 250                 255

Gly Lys Pro Asp Asn Lys Glu Phe Ile Ser His Ala Phe Lys Ile Thr
            260                 265                 270

Gly Glu Gly Leu Lys Val Leu Arg Arg Ala Lys Gly Ser Thr Lys Phe
        275                 280                 285

Thr Arg Val Pro Lys Arg Val Tyr Trp Glu Met Leu Ala Thr Asn Leu
    290                 295                 300

Thr Asp Lys Glu Tyr Val Arg Thr Arg Arg Ala Leu Ile Leu Glu Ile
305                 310                 315                 320

Leu Ile Lys Ala Gly Ser Leu Lys Ile Glu Gln Ile Gln Asp Asn Leu
                325                 330                 335

Lys Lys Leu Gly Phe Asp Glu Val Ile Glu Thr Ile Glu Asn Asp Ile
            340                 345                 350

Lys Gly Leu Ile Asn Thr Gly Ile Phe Ile Glu Ile Lys Gly Arg Phe
        355                 360                 365

Tyr Gln Leu Lys Asp His Ile Leu Gln Phe Val Ile Pro Asn Arg Gly
    370                 375                 380

Val Thr Lys Gln Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu
385                 390                 395                 400

Leu Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile
                405                 410                 415
```

Glu Ile Ala Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val
420 425 430

Met Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly
435 440 445

Gly Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile
450 455 460

Asp Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn
465 470 475 480

Leu Pro Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn
485 490 495

Gln Thr Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr
500 505 510

Pro Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe
515 520 525

Lys Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn
530 535 540

Cys Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu
545 550 555 560

Met Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe
565 570 575

Asn Asn Gly Glu Ile Asn Phe
580

<210> 369

<211> 86
<212> PRT
<213> Artificial sequence

<220>
<223> Im9

<400> 369
Met Glu Leu Lys His Ser Ile Ser Asp Tyr Thr Glu Ala Glu Phe Leu
1 5 10 15

Gln Leu Val Thr Thr Ile Cys Asn Ala Asp Thr Ser Ser Glu Glu Glu
20 25 30

Leu Val Lys Leu Val Thr His Phe Glu Glu Met Thr Glu His Pro Ser
35 40 45

Gly Ser Asp Leu Ile Tyr Tyr Pro Lys Glu Gly Asp Asp Ser Pro
50 55 60

Ser Gly Ile Val Asn Thr Val Lys Gln Trp Arg Ala Ala Asn Gly Lys
65 70 75 80

Ser Gly Phe Lys Gln Gly
85

<210> 370

<211> 1076
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS9082

<400> 370
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
20 25 30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
35 40 45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
50 55 60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65 70 75 80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
85 90 95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
100 105 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
115 120 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
130 135 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145 150 155 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
165 170 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
180 185 190

```
His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
        195                 200                 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
    210                 215                 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225                 230                 235                 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
                245                 250                 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
            260                 265                 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
        275                 280                 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
    290                 295                 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305                 310                 315                 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
                325                 330                 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            340                 345                 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        355                 360                 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
    370                 375                 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385                 390                 395                 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                405                 410                 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            420                 425                 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
        435                 440                 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
    450                 455                 460
```

```
Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465                 470                 475                 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                485                 490                 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
            500                 505                 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        515                 520                 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    530                 535                 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545                 550                 555                 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
                565                 570                 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            580                 585                 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
        595                 600                 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
    610                 615                 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625                 630                 635                 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
                645                 650                 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
            660                 665                 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
        675                 680                 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
    690                 695                 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705                 710                 715                 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                725                 730                 735
```

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
740 745 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
755 760 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
770 775 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785 790 795 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
805 810 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
820 825 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
835 840 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
850 855 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865 870 875 880

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
885 890 895

Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala
900 905 910

Cys Leu Gly Gly Asp Pro Ser Gly Gly Ser Gly Ser Ala Thr Ser Thr
915 920 925

Lys Lys Leu His Lys Glu Pro Ala Thr Leu Ile Lys Ala Ile Asp Gly
930 935 940

Asp Thr Val Lys Leu Met Tyr Lys Gly Gln Pro Met Thr Phe Arg Leu
945 950 955 960

Leu Leu Val Asp Thr Pro Glu Thr Lys His Pro Lys Lys Gly Val Glu
965 970 975

Lys Tyr Gly Pro Glu Ala Ser Ala Phe Thr Lys Lys Met Val Glu Asn
980 985 990

Ala Lys Lys Ile Glu Val Glu Phe Asp Lys Gly Gln Arg Thr Asp Lys
995 1000 1005

Tyr Gly Arg Gly Leu Ala Tyr Ile Tyr Ala Asp Gly Lys Met Val Asn
1010 1015 1020

Glu Ala Leu Val Arg Gln Gly Leu Ala Lys Val Ala Tyr Val Tyr Lys
1025 1030 1035 1040

Pro Asn Asn Thr His Glu Gln His Leu Arg Lys Ser Glu Ala Gln Ala
1045 1050 1055

Lys Lys Glu Lys Leu Asn Ile Trp Ser Glu Asp Asn Ala Asp Ser Gly
1060 1065 1070

Gln Ala Ala Asp
1075

<210> 371

<211> 1088
<212> PRT
<213> Artificial sequence

<220>
<223> pCLS9081

<400> 371
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
20 25 30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
35 40 45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
50 55 60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65 70 75 80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
85 90 95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
100 105 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
115 120 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
130 135 140

```
Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145                 150                 155                 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
                165                 170                 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
            180                 185                 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
        195                 200                 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
    210                 215                 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225                 230                 235                 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
                245                 250                 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
            260                 265                 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
        275                 280                 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
    290                 295                 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305                 310                 315                 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
                325                 330                 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            340                 345                 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        355                 360                 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
    370                 375                 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385                 390                 395                 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                405                 410                 415
```

```
Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            420                 425                 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
        435                 440                 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
    450                 455                 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465                 470                 475                 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                485                 490                 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
            500                 505                 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        515                 520                 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    530                 535                 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545                 550                 555                 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
                565                 570                 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            580                 585                 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
        595                 600                 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
    610                 615                 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625                 630                 635                 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
                645                 650                 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
            660                 665                 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
        675                 680                 685
```

```
Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
    690                 695                 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705                 710                 715                 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                725                 730                 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
            740                 745                 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        755                 760                 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
    770                 775                 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785                 790                 795                 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                805                 810                 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            820                 825                 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
        835                 840                 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    850                 855                 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865                 870                 875                 880

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
                885                 890                 895

Asp Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala
            900                 905                 910

Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Gly
        915                 920                 925

Asp Pro Ser Gly Gly Ser Gly Ser Ala Thr Ser Thr Lys Lys Leu His
    930                 935                 940

Lys Glu Pro Ala Thr Leu Ile Lys Ala Ile Asp Gly Asp Thr Val Lys
945                 950                 955                 960
```

```
Leu Met Tyr Lys Gly Gln Pro Met Thr Phe Arg Leu Leu Leu Val Asp
                965                 970                 975

Thr Pro Glu Thr Lys His Pro Lys Lys Gly Val Glu Lys Tyr Gly Pro
            980                 985                 990

Glu Ala Ser Ala Phe Thr Lys Lys Met Val Glu Asn Ala Lys Lys Ile
        995                 1000                1005

Glu Val Glu Phe Asp Lys Gly Gln Arg Thr Asp Lys Tyr Gly Arg Gly
    1010                1015                1020

Leu Ala Tyr Ile Tyr Ala Asp Gly Lys Met Val Asn Glu Ala Leu Val
1025                1030                1035                1040

Arg Gln Gly Leu Ala Lys Val Ala Tyr Val Tyr Lys Pro Asn Asn Thr
                1045                1050                1055

His Glu Gln His Leu Arg Lys Ser Glu Ala Gln Ala Lys Lys Glu Lys
            1060                1065                1070

Leu Asn Ile Trp Ser Glu Asp Asn Ala Asp Ser Gly Gln Ala Ala Asp
        1075                1080                1085
```

<210> 372

<211> 15
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL1

<400> 372
```
Ser Gly Gly Ser Gly Ser Asn Val Gly Ser Gly Ser Gly Ser Gly
1               5                   10                  15
```

<210> 373

<211> 20
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL2

<400> 373
```
Ser Gly Gly Ser Gly Ser Leu Thr Thr Asn Leu Thr Ala Phe Ser Gly
1               5                   10                  15

Ser Gly Ser Gly
            20
```

<210> 374

<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL3

<400> 374
Ser Gly Gly Ser Gly Ser Lys Arg Arg Thr Thr Ile Ser Ile Ala Ala
1 5 10 15

Ser Gly Ser Gly Ser Gly
20

<210> 375

<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL4

<400> 375
Ser Gly Gly Ser Gly Ser Val Gly Arg Gln Pro Ser Gly Ser Gly Ser
1 5 10 15

Gly

<210> 376

<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL5

<400> 376
Ser Gly Gly Ser Gly Ser Tyr Thr Arg Leu Pro Glu Arg Ser Glu Leu
1 5 10 15

Pro Ala Glu Ile Ser Gly Ser Gly Ser Gly
20 25

<210> 377

<211> 38
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL6

<400> 377
Ser Gly Gly Ser Gly Ser Ile Val Val Leu Asn Arg Ala Glu Thr Pro
1 5 10 15

Leu Pro Leu Asp Pro Thr Gly Lys Val Lys Ala Glu Leu Asp Thr Arg
20 25 30

Ser Gly Ser Gly Ser Gly
35

<210> 378

<211> 21
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL7

<400> 378
Ser Gly Gly Ser Gly Ser Thr Ala Thr Pro Pro Gly Ser Val Thr Ser
1 5 10 15

Gly Ser Gly Ser Gly
20

<210> 379

<211> 21
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL8

<400> 379
Ser Gly Gly Ser Gly Ser Leu Asp Asn Phe Ile Asn Arg Pro Val Ser
1 5 10 15

Gly Ser Gly Ser Gly
20

<210> 380

<211> 21
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL9

<400> 380
Ser Gly Gly Ser Gly Ser Val Ser Ser Ala Lys Thr Thr Ala Pro Ser
1 5 10 15

Gly Ser Gly Ser Gly
20

<210> 381

<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL10

<400> 381
Ser Gly Gly Ser Gly Ser Asp Ser Lys Ala Pro Asn Ala Ser Asn Leu
1 5 10 15

Ser Gly Ser Gly Ser Gly
20

<210> 382

<211> 23
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL11

<400> 382
Ser Gly Gly Ser Gly Ser Pro Val Lys Met Phe Asp Arg His Ser Ser
1 5 10 15

Leu Ser Gly Ser Gly Ser Gly
20

<210> 383

<211> 23
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL12

<400> 383
Ser Gly Gly Ser Gly Ser Ala Pro Ala Glu Thr Lys Ala Glu Pro Met
1 5 10 15

Thr Ser Gly Ser Gly Ser Gly
20

<210> 384

<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL13

<400> 384
Ser Gly Gly Ser Gly Ser Val Ser Thr Asp Ser Thr Pro Val Thr Asn
1 5 10 15

Gln Lys Ser Ser Ser Gly Ser Gly Ser Gly
20 25

<210> 385

<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL14

<400> 385
Ser Gly Gly Ser Gly Ser Asp Ser Val Ile Ser Gly Ser Gly Ser Gly
1 5 10 15

<210> 386

<211> 33
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL15

<400> 386
Ser Gly Gly Ser Gly Ser Ala Arg Phe Thr Leu Ala Val Gly Asp Asn
1 5 10 15

Arg Val Leu Asp Met Ala Ser Thr Tyr Phe Asp Ser Gly Ser Gly Ser
20 25 30

Gly

<210> 387

<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL16

<400> 387
Ser Gly Gly Ser Gly Ser Leu Gln Glu Asn Leu Ser Gly Ser Gly Ser
1 5 10 15

Gly

<210> 388

<211> 19
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL17

<400> 388
Ser Gly Gly Ser Gly Ser Gly Tyr Thr Tyr Ile Asn Pro Ser Gly Ser
1 5 10 15

Gly Ser Gly

<210> 389

<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL18

<400> 389
Ser Gly Gly Ser Gly Ser Tyr Lys Leu Pro Ala Val Thr Thr Met Lys
1 5 10 15

Val Arg Pro Ala Ser Gly Ser Gly Ser Gly
20 25

<210> 390

<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL19

<400> 390
Ser Gly Gly Ser Gly Ser Leu Glu Gly Ser Ser Gly Ser Gly Ser Gly
1 5 10 15

<210> 391

<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL20

<400> 391
Ser Gly Gly Ser Gly Ser Ile Val Glu Ala Ser Gly Ser Gly Ser Gly
1 5 10 15

<210> 392

<211> 18
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL21

<400> 392
Ser Gly Gly Ser Gly Ser Gln Thr Tyr Gln Pro Ala Ser Gly Ser Gly
1 5 10 15

Ser Gly

<210> 393

<211> 27
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL22

<400> 393
Ser Gly Gly Ser Gly Ser Ile Ala Arg Thr Asp Leu Lys Lys Asn Arg
1 5 10 15

Asp Tyr Pro Leu Ala Ser Gly Ser Gly Ser Gly
20 25

<210> 394

<211> 18
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL23

<400> 394
Ser Gly Gly Ser Gly Ser Leu Pro Glu Glu Lys Gly Ser Gly Ser Gly
1 5 10 15

Ser Gly

<210> 395

<211> 16
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL24

<400> 395
Ser Gly Gly Ser Gly Ser Tyr Thr Ser Thr Ser Gly Ser Gly Ser Gly
1 5 10 15

<210> 396

<211> 20
<212> PRT

<213> Artificial sequence

<220>
<223> Linker_TAL25

<400> 396
Ser Gly Gly Ser Gly Ser Ser Arg Pro Ser Glu Ser Glu Gly Ser Gly
1 5 10 15

Ser Gly Ser Gly
20

<210> 397

<211> 17
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL26

<400> 397
Ser Gly Gly Ser Gly Ser Leu Gly Asn Ser Leu Ser Gly Ser Gly Ser
1 5 10 15

Gly

<210> 398

<211> 19
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL27

<400> 398
Ser Gly Gly Ser Gly Ser Leu Thr Lys Tyr Lys Ser Ser Ser Gly Ser
1 5 10 15

Gly Ser Gly

<210> 399

<211> 33
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL28

<400> 399
Ser Gly Gly Ser Gly Ser Thr Glu Glu Pro Gly Ala Pro Leu Thr Thr
1 5 10 15

Pro Pro Thr Leu His Gly Asn Gln Ala Arg Ala Ser Gly Ser Gly Ser
20 25 30

Gly

<210> 400

<211> 18
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL29

<400> 400
Ser Gly Gly Ser Gly Ser Phe Ser His Ser Thr Thr Ser Gly Ser Gly
1 5 10 15

Ser Gly

<210> 401

<211> 20
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL30

<400> 401
Ser Gly Gly Ser Gly Ser Pro Glu Leu Lys Gln Lys Ser Ser Ser Gly
1 5 10 15

Ser Gly Ser Gly
20

<210> 402

<211> 40
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL31

<400> 402
Ser Gly Gly Ser Gly Ser Ile Leu Asn Lys Glu Lys Lys Ala Val Ser
1 5 10 15

Pro Leu Leu Leu Thr Thr Thr Asn Ser Ser Glu Gly Leu Ser Met Gly
20 25 30

Asn Tyr Ser Gly Ser Gly Ser Gly
35 40

<210> 403

<211> 31
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL32

<400> 403
Glu Leu Ala Glu Phe His Ala Arg Tyr Ala Asp Leu Leu Leu Arg Asp
1 5 10 15

Leu Arg Glu Arg Pro Val Ser Leu Val Arg Gly Pro Asp Ser Gly
20 25 30

<210> 404

<211> 31
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL33

<400> 404
Glu Leu Ala Glu Phe His Ala Arg Pro Asp Pro Leu Leu Leu Arg Asp
1 5 10 15

Leu Arg Glu Arg Pro Val Ser Leu Val Arg Gly Leu Gly Ser Gly
20 25 30

<210> 405

<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL34

<400> 405
Glu Leu Ala Glu Phe His Ala Arg Tyr Ala Asp Leu Leu Leu Arg Asp
1 5 10 15

Leu Arg Glu Arg Ser Gly Ser Gly Ser Gly
20 25

<210> 406

<211> 31
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL35

<400> 406
Asp Ile Phe Asp Tyr Tyr Ala Gly Val Ala Glu Val Met Leu Gly His
1 5 10 15

Ile Ala Gly Arg Pro Ala Thr Arg Lys Arg Trp Pro Asn Ser Gly
20 25 30

<210> 407

<211> 31
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL36

<400> 407
Asp Ile Phe Asp Tyr Tyr Ala Gly Pro Asp Pro Val Met Leu Gly His
1 5 10 15

Ile Ala Gly Arg Pro Ala Thr Arg Lys Arg Trp Leu Gly Ser Gly
20 25 30

<210> 408

<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Linker_TAL37

<400> 408
Asp Ile Phe Asp Tyr Tyr Ala Gly Val Ala Glu Val Met Leu Gly His
1 5 10 15

Ile Ala Gly Arg Ser Gly Ser Gly Ser Gly
20 25

<210> 409

<211> 37
<212> PRT
<213> Artificial sequence

<220>
<223> Linker A

<400> 409
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Val Ser Phe
1 5 10 15

Gln Lys Leu Lys Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala
20 25 30

Val Lys Lys Gly Leu
35

<210> 410

<211> 37
<212> PRT
<213> Artificial sequence

<220>
<223> Linker B

<400> 410
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Ala Val Ser Ala
1 5 10 15

Gln Lys Ala Lys Ala Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala
20 25 30

Val Lys Lys Gly Leu
35

<210> 411

<211> 37
<212> PRT
<213> Artificial sequence

<220>
<223> Linker C

<400> 411
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Val Val Thr Phe
1 5 10 15

His Lys Leu Lys Leu Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp Ala
20 25 30

Val Lys Lys Gly Leu
35

<210> 412

<211> 44
<212> PRT
<213> Artificial sequence

<220>
<223> Linker D

<400> 412
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Gln Ser Leu Ala
1 5 10 15

Gln Glu Leu Ser Leu Asn Glu Ser Gln Ile Lys Ile Ala Cys Leu Gly
20 25 30

Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu
35 40

<210> 413

<211> 40
<212> PRT
<213> Artificial sequence

<220>
<223> Linker E

<400> 413
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Gln Leu Pro
1 5 10 15

Pro Leu Glu Arg Leu Thr Leu Asp Ala Cys Leu Gly Gly Arg Pro Ala
20 25 30

Leu Asp Ala Val Lys Lys Gly Leu
35 40

<210> 414

<211> 38
<212> PRT
<213> Artificial sequence

<220>
<223> Linker F

<400> 414
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Ile His Lys Lys
1 5 10 15

Phe Ser Ser Ile Gln Met Ala Cys Leu Gly Gly Arg Pro Ala Leu Asp
20 25 30

Ala Val Lys Lys Gly Leu
35

<210> 415

<211> 40
<212> PRT
<213> Artificial sequence

<220>
<223> Linker G

<400> 415
Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Ala Ala Ala Ala
1 5 10 15

Thr Asn Asp His Ala Val Ala Ala Ala Cys Leu Gly Gly Arg Pro Ala
20 25 30

Leu Asp Ala Val Lys Lys Gly Leu
35 40

<210> 416

<211> 1044
<212> PRT
<213> Artificial sequence

<220>
<223> Linker385_ColE7

<400> 416
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
20 25 30

```
Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
        35                  40                  45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
    50                  55                  60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65                  70                  75                  80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
                85                  90                  95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
            100                 105                 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
        115                 120                 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
    130                 135                 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145                 150                 155                 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
                165                 170                 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
            180                 185                 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
        195                 200                 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
    210                 215                 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225                 230                 235                 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
                245                 250                 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
            260                 265                 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
        275                 280                 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
    290                 295                 300
```

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305 310 315 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
325 330 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
340 345 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
355 360 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
370 375 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385 390 395 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
405 410 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
420 425 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
435 440 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
450 455 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465 470 475 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
485 490 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
500 505 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
515 520 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
530 535 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545 550 555 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
565 570 575

```
Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            580                 585                 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
        595                 600                 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
    610                 615                 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625                 630                 635                 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
                645                 650                 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
            660                 665                 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
        675                 680                 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
    690                 695                 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705                 710                 715                 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                725                 730                 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
            740                 745                 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        755                 760                 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
    770                 775                 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785                 790                 795                 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                805                 810                 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            820                 825                 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
        835                 840                 845
```

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
850 855 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865 870 875 880

Gly Gly Arg Pro Ala Leu Glu Ser Gly Gly Ser Gly Ser Asp Ser Val
885 890 895

Ile Ser Gly Ser Gly Ser Gly Asp Pro Ser Gly Gly Ser Gly Ser Lys
900 905 910

Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys Gly Lys Pro Val Asn Asn
915 920 925

Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu Gly Ser Pro Val Pro Asp
930 935 940

Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu Phe Lys Ser Phe Asp Asp
945 950 955 960

Phe Arg Lys Lys Phe Trp Glu Glu Val Ser Lys Asp Pro Glu Leu Ser
965 970 975

Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg Met Lys Val Gly Lys Ala
980 985 990

Pro Lys Thr Arg Thr Gln Asp Val Ser Gly Lys Arg Thr Ser Phe Glu
995 1000 1005

Leu His His Glu Lys Pro Ile Ser Gln Asn Gly Gly Val Tyr Asp Met
1010 1015 1020

Asp Asn Ile Ser Val Val Thr Pro Lys Arg His Ile Asp Ile His Arg
1025 1030 1035 1040

Gly Lys Ala Ala

<210> 417

<211> 1047
<212> PRT
<213> Artificial sequence

<220>
<223> Linker388_ColE7

<400> 417
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

```
Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
            20                  25                  30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
        35                  40                  45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
    50                  55                  60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65                  70                  75                  80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
                85                  90                  95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
            100                 105                 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
        115                 120                 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
    130                 135                 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145                 150                 155                 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
                165                 170                 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
            180                 185                 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
        195                 200                 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
    210                 215                 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225                 230                 235                 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
                245                 250                 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
            260                 265                 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
        275                 280                 285
```

```
Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
    290                 295                 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305                 310                 315                 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
                325                 330                 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            340                 345                 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        355                 360                 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
    370                 375                 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385                 390                 395                 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                405                 410                 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            420                 425                 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
        435                 440                 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
    450                 455                 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465                 470                 475                 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                485                 490                 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
            500                 505                 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        515                 520                 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    530                 535                 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545                 550                 555                 560
```

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
565 570 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
580 585 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
595 600 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
610 615 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625 630 635 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
645 650 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
660 665 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
675 680 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
690 695 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705 710 715 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
725 730 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
740 745 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
755 760 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
770 775 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785 790 795 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
805 810 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
820 825 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
835 840 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
850 855 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865 870 875 880

Gly Gly Arg Pro Ala Leu Glu Ser Gly Gly Ser Gly Ser Gly Tyr Thr
885 890 895

Tyr Ile Asn Pro Ser Gly Ser Gly Ser Gly Asp Pro Ser Gly Gly Ser
900 905 910

Gly Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys Gly Lys Pro
915 920 925

Val Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu Gly Ser Pro
930 935 940

Val Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu Phe Lys Ser
945 950 955 960

Phe Asp Asp Phe Arg Lys Lys Phe Trp Glu Glu Val Ser Lys Asp Pro
965 970 975

Glu Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg Met Lys Val
980 985 990

Gly Lys Ala Pro Lys Thr Arg Thr Gln Asp Val Ser Gly Lys Arg Thr
995 1000 1005

Ser Phe Glu Leu His His Glu Lys Pro Ile Ser Gln Asn Gly Gly Val
1010 1015 1020

Tyr Asp Met Asp Asn Ile Ser Val Val Thr Pro Lys Arg His Ile Asp
1025 1030 1035 1040

Ile His Arg Gly Lys Ala Ala
1045

<210> 418

<211> 1046
<212> PRT
<213> Artificial sequence

<220>
<223> Linker394_ColE7

<400> 418
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1 5 10 15

```
Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
            20                  25                  30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
        35                  40                  45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
    50                  55                  60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65                  70                  75                  80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
                85                  90                  95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
            100                 105                 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
        115                 120                 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
    130                 135                 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145                 150                 155                 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
                165                 170                 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
            180                 185                 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
        195                 200                 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
    210                 215                 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225                 230                 235                 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
                245                 250                 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
            260                 265                 270

Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
        275                 280                 285
```

```
Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
    290                 295                 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305                 310                 315                 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
                325                 330                 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            340                 345                 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        355                 360                 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
    370                 375                 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385                 390                 395                 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                405                 410                 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            420                 425                 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
        435                 440                 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
    450                 455                 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465                 470                 475                 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                485                 490                 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
            500                 505                 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        515                 520                 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    530                 535                 540

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545                 550                 555                 560
```

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
565 570 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
580 585 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
595 600 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
610 615 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625 630 635 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
645 650 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
660 665 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
675 680 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
690 695 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705 710 715 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
725 730 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
740 745 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
755 760 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
770 775 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785 790 795 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
805 810 815

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
820 825 830

```
Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
        835                 840                 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    850                 855                 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865                 870                 875                 880

Gly Gly Arg Pro Ala Leu Glu Ser Gly Gly Ser Gly Ser Leu Pro Glu
                885                 890                 895

Glu Lys Gly Ser Gly Ser Gly Ser Gly Asp Pro Ser Gly Gly Ser Gly
            900                 905                 910

Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys Gly Lys Pro Val
        915                 920                 925

Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu Gly Ser Pro Val
    930                 935                 940

Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu Phe Lys Ser Phe
945                 950                 955                 960

Asp Asp Phe Arg Lys Lys Phe Trp Glu Glu Val Ser Lys Asp Pro Glu
                965                 970                 975

Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg Met Lys Val Gly
            980                 985                 990

Lys Ala Pro Lys Thr Arg Thr Gln Asp Val Ser Gly Lys Arg Thr Ser
        995                 1000                1005

Phe Glu Leu His His Glu Lys Pro Ile Ser Gln Asn Gly Gly Val Tyr
    1010                1015                1020

Asp Met Asp Asn Ile Ser Val Val Thr Pro Lys Arg His Ile Asp Ile
1025                1030                1035                1040

His Arg Gly Lys Ala Ala
                1045
```

<210> 419

<211> 1046
<212> PRT
<213> Artificial sequence

<220>
<223> Linker400_ColE7

<400> 419

```
Met Ala Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu
1               5                  10                  15

Leu Pro Gly Pro Gln Pro Asp Gly Val Gln Pro Thr Ala Asp Arg Gly
            20                  25                  30

Val Ser Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg
        35                  40                  45

Thr Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala
    50                  55                  60

Phe Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser
65                  70                  75                  80

Leu Phe Asn Thr Ser Leu Phe Asp Ser Leu Pro Pro Phe Gly Ala His
                85                  90                  95

His Thr Glu Ala Ala Thr Gly Glu Trp Asp Glu Val Gln Ser Gly Leu
            100                 105                 110

Arg Ala Ala Asp Ala Pro Pro Pro Thr Met Arg Val Ala Val Thr Ala
        115                 120                 125

Ala Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln
    130                 135                 140

Pro Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly
145                 150                 155                 160

Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr
                165                 170                 175

Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala
            180                 185                 190

His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala
        195                 200                 205

Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu
    210                 215                 220

Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu
225                 230                 235                 240

Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu
                245                 250                 255

Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala
            260                 265                 270
```

```
Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu
        275                 280                 285

Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
    290                 295                 300

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
305                 310                 315                 320

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
                325                 330                 335

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            340                 345                 350

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        355                 360                 365

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
    370                 375                 380

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
385                 390                 395                 400

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                405                 410                 415

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            420                 425                 430

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
        435                 440                 445

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
    450                 455                 460

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
465                 470                 475                 480

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                485                 490                 495

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
            500                 505                 510

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        515                 520                 525

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    530                 535                 540
```

```
Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
545                 550                 555                 560

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
                565                 570                 575

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            580                 585                 590

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
        595                 600                 605

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
    610                 615                 620

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
625                 630                 635                 640

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
                645                 650                 655

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
            660                 665                 670

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
        675                 680                 685

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
    690                 695                 700

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
705                 710                 715                 720

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                725                 730                 735

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
            740                 745                 750

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        755                 760                 765

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
    770                 775                 780

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
785                 790                 795                 800

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                805                 810                 815
```

```
Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            820                 825                 830

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
        835                 840                 845

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    850                 855                 860

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
865                 870                 875                 880

Gly Gly Arg Pro Ala Leu Glu Ser Gly Gly Ser Gly Ser Phe Ser His
                885                 890                 895

Ser Thr Thr Ser Gly Ser Gly Ser Gly Asp Pro Ser Gly Gly Ser Gly
            900                 905                 910

Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys Gly Lys Pro Val
        915                 920                 925

Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu Gly Ser Pro Val
    930                 935                 940

Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu Phe Lys Ser Phe
945                 950                 955                 960

Asp Asp Phe Arg Lys Lys Phe Trp Glu Glu Val Ser Lys Asp Pro Glu
                965                 970                 975

Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg Met Lys Val Gly
            980                 985                 990

Lys Ala Pro Lys Thr Arg Thr Gln Asp Val Ser Gly Lys Arg Thr Ser
        995                 1000                1005

Phe Glu Leu His His Glu Lys Pro Ile Ser Gln Asn Gly Gly Val Tyr
    1010                1015                1020

Asp Met Asp Asn Ile Ser Val Val Thr Pro Lys Arg His Ile Asp Ile
1025                1030                1035                1040

His Arg Gly Lys Ala Ala
                1045
```

<210> 420

<211> 365
<212> PRT
<213> Artificial sequence

<220>
<223> cT11_TevD01

<400> 420
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
100 105 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
115 120 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly
130 135 140

Ile Arg Ser Arg Ala Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro
145 150 155 160

Asp Pro Gly Ser Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn
165 170 175

Asn Lys Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys
180 185 190

Arg His Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu
195 200 205

Gln Arg Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu
210 215 220

Glu Glu Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe
225 230 235 240

Trp Ile Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp
245 250 255

Ala Thr Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile
260 265 270

Ile Lys Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly
275 280 285

Pro Asp Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly
290 295 300

Arg Trp Asn Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile
305 310 315 320

Gln Thr Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu
325 330 335

Asn Asn Ser Phe Phe Asn His Lys His Ser Asp Ile Thr Lys Ser Lys
340 345 350

Ile Ser Gly Lys Met Lys Gly Gln Gly Pro Ser Ala Asp
355 360 365

<210> 421

<211> 352
<212> PRT
<213> Artificial sequence

<220>
<223> cT11_TevD02

<400> 421
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
100 105 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
115 120 125

```
Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly
    130                 135                 140

Ile Arg Ser Arg Ala Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro
145                 150                 155                 160

Asp Pro Gly Ser Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn
                165                 170                 175

Asn Lys Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys
            180                 185                 190

Arg His Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu
        195                 200                 205

Gln Arg Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu
    210                 215                 220

Glu Glu Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe
225                 230                 235                 240

Trp Ile Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp
                245                 250                 255

Ala Thr Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile
            260                 265                 270

Ile Lys Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly
        275                 280                 285

Pro Asp Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly
    290                 295                 300

Arg Trp Asn Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile
305                 310                 315                 320

Gln Thr Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu
                325                 330                 335

Asn Asn Ser Phe Phe Asn His Lys His Ser Gln Gly Pro Ser Ala Asp
            340                 345                 350
```

<210> 422

<211> 292
<212> PRT
<213> Artificial sequence

<220>
<223> cT11_TevD05

<400> 422
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
100 105 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
115 120 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly
130 135 140

Ile Arg Ser Arg Ala Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro
145 150 155 160

Asp Pro Gly Ser Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn
165 170 175

Asn Lys Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys
180 185 190

Arg His Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu
195 200 205

Gln Arg Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu
210 215 220

Glu Glu Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe
225 230 235 240

Trp Ile Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp
245 250 255

Ala Thr Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile
260 265 270

Ile Lys Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Gln Gly
        275   280   285

Pro Ser Ala Asp
    290

<210> 423

<211> 949
<212> PRT
<213> Artificial sequence

<220>
<223> cT11Avr_TevD01

<400> 423
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Glu Lys
1   5   10   15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
            20   25   30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
        35   40   45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
    50   55   60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65   70   75   80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
                85   90   95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
            100   105   110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115   120   125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val
    130   135   140

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145   150   155   160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
                165   170   175

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
            180   185   190

```
Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        195                 200                 205

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
    210                 215                 220

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225                 230                 235                 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                245                 250                 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            260                 265                 270

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
        275                 280                 285

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
    290                 295                 300

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile
305                 310                 315                 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu
                325                 330                 335

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
            340                 345                 350

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
        355                 360                 365

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
    370                 375                 380

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
385                 390                 395                 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
                405                 410                 415

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
            420                 425                 430

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
        435                 440                 445

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
    450                 455                 460
```

```
Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465                 470                 475                 480

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
                485                 490                 495

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
            500                 505                 510

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
        515                 520                 525

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
    530                 535                 540

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
545                 550                 555                 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
                565                 570                 575

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
            580                 585                 590

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
        595                 600                 605

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
    610                 615                 620

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
625                 630                 635                 640

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
                645                 650                 655

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
            660                 665                 670

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
        675                 680                 685

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
    690                 695                 700

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
705                 710                 715                 720

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser
                725                 730                 735
```

Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Gly Ser Lys Ser Gly Ile
740 745 750

Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val Tyr Val Gly Ser Ala
755 760 765

Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe Lys Asp Leu Glu Lys
770 775 780

Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser Phe Asn Lys His Gly
785 790 795 800

Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile Pro Tyr Glu Lys Asp
805 810 815

Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys Glu Leu Asn Ser Lys
820 825 830

Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe Gly Asp Thr Cys Ser
835 840 845

Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys Arg Ser Glu Thr Val
850 855 860

Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly Arg Lys Ala Leu Tyr
865 870 875 880

Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp Asn Pro Glu Thr His Lys
885 890 895

Phe Cys Lys Cys Gly Val Arg Ile Gln Thr Ser Ala Tyr Thr Cys Ser
900 905 910

Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn Ser Phe Phe Asn His Lys
915 920 925

His Ser Asp Ile Thr Lys Ser Lys Ile Ser Gly Lys Met Lys Gly Gln
930 935 940

Gly Pro Ser Ala Asp
945

<210> 424

<211> 936
<212> PRT
<213> Artificial sequence

<220>
<223> cT11Avr_TevD02

<400> 424
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

```
Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
            20                  25                  30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
        35                  40                  45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
    50                  55                  60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65                  70                  75                  80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
                85                  90                  95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
            100                 105                 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val
    130                 135                 140

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145                 150                 155                 160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
                165                 170                 175

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
            180                 185                 190

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        195                 200                 205

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
    210                 215                 220

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225                 230                 235                 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                245                 250                 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            260                 265                 270

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
        275                 280                 285
```

```
Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
    290                 295                 300

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile
305                 310                 315                 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu
                325                 330                 335

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
            340                 345                 350

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
        355                 360                 365

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
    370                 375                 380

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
385                 390                 395                 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
                405                 410                 415

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
            420                 425                 430

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
        435                 440                 445

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
    450                 455                 460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465                 470                 475                 480

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
                485                 490                 495

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
            500                 505                 510

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
        515                 520                 525

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
    530                 535                 540

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
545                 550                 555                 560
```

```
Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
                565                 570                 575

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
            580                 585                 590

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
        595                 600                 605

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
    610                 615                 620

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
625                 630                 635                 640

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
                645                 650                 655

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
            660                 665                 670

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
        675                 680                 685

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
    690                 695                 700

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
705                 710                 715                 720

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser
                725                 730                 735

Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Gly Ser Lys Ser Gly Ile
            740                 745                 750

Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val Tyr Val Gly Ser Ala
        755                 760                 765

Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe Lys Asp Leu Glu Lys
    770                 775                 780

Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser Phe Asn Lys His Gly
785                 790                 795                 800

Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile Pro Tyr Glu Lys Asp
                805                 810                 815

Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys Glu Leu Asn Ser Lys
            820                 825                 830
```

Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe Gly Asp Thr Cys Ser
835 840 845

Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys Arg Ser Glu Thr Val
850 855 860

Lys Ala Lys Met Leu Lys Leu Gly Pro Asp Gly Arg Lys Ala Leu Tyr
865 870 875 880

Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp Asn Pro Glu Thr His Lys
885 890 895

Phe Cys Lys Cys Gly Val Arg Ile Gln Thr Ser Ala Tyr Thr Cys Ser
900 905 910

Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn Ser Phe Phe Asn His Lys
915 920 925

His Ser Gln Gly Pro Ser Ala Asp
930 935

<210> 425

<211> 876
<212> PRT
<213> Artificial sequence

<220>
<223> cT11Avr_TevD05

<400> 425
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
100 105 110

```
Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val
    130                 135                 140

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145                 150                 155                 160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
                165                 170                 175

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
            180                 185                 190

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        195                 200                 205

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
    210                 215                 220

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225                 230                 235                 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                245                 250                 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            260                 265                 270

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
        275                 280                 285

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
    290                 295                 300

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile
305                 310                 315                 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu
                325                 330                 335

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
            340                 345                 350

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
        355                 360                 365

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
    370                 375                 380
```

```
Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
385                 390                 395                 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
                405                 410                 415

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
            420                 425                 430

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
        435                 440                 445

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
    450                 455                 460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465                 470                 475                 480

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
                485                 490                 495

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
            500                 505                 510

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
        515                 520                 525

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
    530                 535                 540

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
545                 550                 555                 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
                565                 570                 575

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
            580                 585                 590

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
        595                 600                 605

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
    610                 615                 620

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
625                 630                 635                 640

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
                645                 650                 655
```

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
660 665 670

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
675 680 685

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
690 695 700

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
705 710 715 720

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser
725 730 735

Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Gly Ser Lys Ser Gly Ile
740 745 750

Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys Val Tyr Val Gly Ser Ala
755 760 765

Lys Asp Phe Glu Lys Arg Trp Lys Arg His Phe Lys Asp Leu Glu Lys
770 775 780

Gly Cys His Ser Ser Ile Lys Leu Gln Arg Ser Phe Asn Lys His Gly
785 790 795 800

Asn Val Phe Glu Cys Ser Ile Leu Glu Glu Ile Pro Tyr Glu Lys Asp
805 810 815

Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile Lys Glu Leu Asn Ser Lys
820 825 830

Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr Phe Gly Asp Thr Cys Ser
835 840 845

Thr His Pro Leu Lys Glu Glu Ile Ile Lys Lys Arg Ser Glu Thr Val
850 855 860

Lys Ala Lys Met Leu Lys Gln Gly Pro Ser Ala Asp
865 870 875

<210> 426

<211> 368
<212> PRT
<213> Artificial sequence

<220>
<223> TevW01_cT11

<400> 426
Met Ala Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys
1 5 10 15

```
Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His
            20                  25                  30

Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg
        35                  40                  45

Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu
    50                  55                  60

Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile
65                  70                  75                  80

Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr
                85                  90                  95

Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys
            100                 105                 110

Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp
        115                 120                 125

Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp
    130                 135                 140

Asn Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr
145                 150                 155                 160

Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn
                165                 170                 175

Ser Phe Phe Asn His Lys His Ser Asp Ile Thr Lys Ser Lys Ile Ser
            180                 185                 190

Gly Lys Met Lys Gly Lys Lys Pro Ser Asn Ser Gly Val Asp Leu Arg
        195                 200                 205

Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val
    210                 215                 220

Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe
225                 230                 235                 240

Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly
                245                 250                 255

Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala
            260                 265                 270

Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg
        275                 280                 285
```

```
Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro
    290                 295                 300

Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly
305                 310                 315                 320

Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly
                325                 330                 335

Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly Ile Arg Ser Arg Ala
            340                 345                 350

Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ser Ala Asp
        355                 360                 365
```

<210> 427

<211> 366
<212> PRT
<213> Artificial sequence

<220>
<223> TevD01_cT11

<400> 427

```
Met Ala Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys
1               5                   10                  15

Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His
            20                  25                  30

Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg
        35                  40                  45

Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu
    50                  55                  60

Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile
65                  70                  75                  80

Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr
                85                  90                  95

Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys
            100                 105                 110

Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp
        115                 120                 125
```

```
Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp
    130                 135                 140

Asn Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr
145                 150                 155                 160

Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn
                165                 170                 175

Ser Phe Phe Asn His Lys His Ser Asp Ile Thr Lys Ser Lys Ile Ser
            180                 185                 190

Gly Lys Met Lys Gly Gln Gly Pro Ser Gly Val Asp Leu Arg Thr Leu
        195                 200                 205

Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser
    210                 215                 220

Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His
225                 230                 235                 240

Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val
                245                 250                 255

Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His
            260                 265                 270

Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu
        275                 280                 285

Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln
    290                 295                 300

Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr
305                 310                 315                 320

Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro
                325                 330                 335

Leu Asn Leu Thr Gly Asp Ala Arg Gly Ile Arg Ser Arg Ala Ser Arg
            340                 345                 350

Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ser Ala Asp
        355                 360                 365
```

<210> 428

<211> 353
<212> PRT
<213> Artificial sequence

<220>
<223> TevD02_cT11

<400> 428
Met Ala Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys
1 5 10 15

Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His
20 25 30

Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg
35 40 45

Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu
50 55 60

Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile
65 70 75 80

Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr
85 90 95

Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys
100 105 110

Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp
115 120 125

Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp
130 135 140

Asn Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr
145 150 155 160

Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn
165 170 175

Ser Phe Phe Asn His Lys His Ser Gln Gly Pro Ser Gly Val Asp Leu
180 185 190

Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys
195 200 205

Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly
210 215 220

Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu
225 230 235 240

Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu
245 250 255

Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala
260 265 270

Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro
275 280 285

Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly
290 295 300

Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr
305 310 315 320

Gly Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly Ile Arg Ser Arg
325 330 335

Ala Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ser Ala
340 345 350

Asp

<210> 429

<211> 293
<212> PRT
<213> Artificial sequence

<220>
<223> TevD05_cT11

<400> 429
Met Ala Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys
1 5 10 15

Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His
20 25 30

Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg
35 40 45

Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu
50 55 60

Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile
65 70 75 80

Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr
85 90 95

Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys
100 105 110

Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Gln Gly Pro Ser
115 120 125

Gly Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
130 135 140

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
145 150 155 160

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
165 170 175

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
180 185 190

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
195 200 205

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
210 215 220

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
225 230 235 240

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
245 250 255

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly
260 265 270

Ile Arg Ser Arg Ala Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro
275 280 285

Asp Pro Ser Ala Asp
290

<210> 430

<211> 952
<212> PRT
<213> Artificial sequence

<220>
<223> TevW01_cT11Avr

<400> 430
Met Ala Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys
1 5 10 15

Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His
20 25 30

Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg
35 40 45

Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu
50 55 60

```
Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile
65                  70                  75                  80

Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr
                85                  90                  95

Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys
            100                 105                 110

Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp
        115                 120                 125

Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp
    130                 135                 140

Asn Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr
145                 150                 155                 160

Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn
                165                 170                 175

Ser Phe Phe Asn His Lys His Ser Asp Ile Thr Lys Ser Lys Ile Ser
            180                 185                 190

Gly Lys Met Lys Gly Lys Lys Pro Ser Asn Ser Gly Val Asp Leu Arg
        195                 200                 205

Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val
    210                 215                 220

Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe
225                 230                 235                 240

Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly
                245                 250                 255

Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala
            260                 265                 270

Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg
        275                 280                 285

Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro
    290                 295                 300

Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly
305                 310                 315                 320

Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly
                325                 330                 335
```

```
Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His
            340                 345                 350

Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
        355                 360                 365

Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
    370                 375                 380

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
385                 390                 395                 400

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
                405                 410                 415

Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala
            420                 425                 430

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
        435                 440                 445

Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val
    450                 455                 460

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
465                 470                 475                 480

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu
                485                 490                 495

Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
            500                 505                 510

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala
        515                 520                 525

Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly
    530                 535                 540

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys
545                 550                 555                 560

Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala
                565                 570                 575

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly
            580                 585                 590

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
        595                 600                 605
```

```
Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His
    610                 615                 620

Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
625                 630                 635                 640

Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
                645                 650                 655

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
            660                 665                 670

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
        675                 680                 685

Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala
    690                 695                 700

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
705                 710                 715                 720

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val
                725                 730                 735

Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
            740                 745                 750

Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu
        755                 760                 765

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
    770                 775                 780

Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala
785                 790                 795                 800

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
                805                 810                 815

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
            820                 825                 830

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
        835                 840                 845

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
    850                 855                 860

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
865                 870                 875                 880
```

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His
885 890 895

Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
900 905 910

Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
915 920 925

Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu
930 935 940

Ser Arg Pro Asp Pro Ser Ala Asp
945 950

<210> 431

<211> 950
<212> PRT
<213> Artificial sequence

<220>
<223> TevD01_cT11Avr

<400> 431
Met Ala Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys
1 5 10 15

Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His
20 25 30

Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg
35 40 45

Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu
50 55 60

Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile
65 70 75 80

Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr
85 90 95

Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys
100 105 110

Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp
115 120 125

Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp
130 135 140

```
Asn Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr
145                 150                 155                 160

Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn
                165                 170                 175

Ser Phe Phe Asn His Lys His Ser Asp Ile Thr Lys Ser Lys Ile Ser
            180                 185                 190

Gly Lys Met Lys Gly Gln Gly Pro Ser Gly Val Asp Leu Arg Thr Leu
        195                 200                 205

Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser
    210                 215                 220

Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His
225                 230                 235                 240

Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val
                245                 250                 255

Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His
            260                 265                 270

Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu
        275                 280                 285

Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln
    290                 295                 300

Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr
305                 310                 315                 320

Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro
                325                 330                 335

Leu Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly
            340                 345                 350

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
        355                 360                 365

Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
    370                 375                 380

Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
385                 390                 395                 400

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
                405                 410                 415
```

```
Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu
            420                 425                 430

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
        435                 440                 445

Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
    450                 455                 460

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
465                 470                 475                 480

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val
                485                 490                 495

Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
            500                 505                 510

Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu
        515                 520                 525

Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
    530                 535                 540

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala
545                 550                 555                 560

Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly
                565                 570                 575

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
            580                 585                 590

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
        595                 600                 605

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly
    610                 615                 620

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
625                 630                 635                 640

Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
                645                 650                 655

Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
            660                 665                 670

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
        675                 680                 685
```

```
Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu
    690                 695                 700

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
705                 710                 715                 720

Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala
                725                 730                 735

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
            740                 745                 750

Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val
        755                 760                 765

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
    770                 775                 780

Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu
785                 790                 795                 800

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
                805                 810                 815

Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala
            820                 825                 830

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
        835                 840                 845

Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys
    850                 855                 860

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
865                 870                 875                 880

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly
                885                 890                 895

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
            900                 905                 910

Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
        915                 920                 925

Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg
    930                 935                 940

Pro Asp Pro Ser Ala Asp
945                 950
```

<210> 432

<211> 937
<212> PRT
<213> Artificial sequence

<220>
<223> TevD02_cT11Avr

<400> 432

```
Met Ala Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys
1               5                   10                  15

Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His
            20                  25                  30

Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg
        35                  40                  45

Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu
    50                  55                  60

Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile
65                  70                  75                  80

Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr
                85                  90                  95

Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys
            100                 105                 110

Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp
        115                 120                 125

Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp
    130                 135                 140

Asn Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr
145                 150                 155                 160

Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn
                165                 170                 175

Ser Phe Phe Asn His Lys His Ser Gln Gly Pro Ser Gly Val Asp Leu
            180                 185                 190

Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys
        195                 200                 205

Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly
    210                 215                 220

Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu
225                 230                 235                 240
```

Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu
245 250 255

Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala
260 265 270

Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro
275 280 285

Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly
290 295 300

Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr
305 310 315 320

Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
325 330 335

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
340 345 350

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
355 360 365

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
370 375 380

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
385 390 395 400

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
405 410 415

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
420 425 430

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
435 440 445

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
450 455 460

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
465 470 475 480

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
485 490 495

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
500 505 510

```
Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
        515                 520                 525

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
    530                 535                 540

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
545                 550                 555                 560

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
                565                 570                 575

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            580                 585                 590

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        595                 600                 605

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
    610                 615                 620

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
625                 630                 635                 640

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                645                 650                 655

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            660                 665                 670

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
        675                 680                 685

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
    690                 695                 700

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
705                 710                 715                 720

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                725                 730                 735

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
            740                 745                 750

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        755                 760                 765

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    770                 775                 780
```

```
Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
785                 790                 795                 800

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
                805                 810                 815

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            820                 825                 830

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
        835                 840                 845

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
    850                 855                 860

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
865                 870                 875                 880

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
                885                 890                 895

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
            900                 905                 910

Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln
        915                 920                 925

Leu Ser Arg Pro Asp Pro Ser Ala Asp
    930                 935


<210>  433

<211>  417
<212>  PRT
<213>  Artificial sequence

<220>
<223>  cT11_NucA

<400>  433
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1               5                   10                  15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
            20                  25                  30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
        35                  40                  45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
    50                  55                  60
```

```
Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65                  70                  75                  80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
                85                  90                  95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
            100                 105                 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly
    130                 135                 140

Ile Arg Ser Arg Ala Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro
145                 150                 155                 160

Asp Pro Gly Ser Gln Val Pro Pro Leu Thr Glu Leu Ser Pro Ser Ile
                165                 170                 175

Ser Val His Leu Leu Leu Gly Asn Pro Ser Gly Ala Thr Pro Thr Lys
            180                 185                 190

Leu Thr Pro Asp Asn Tyr Leu Met Val Lys Asn Gln Tyr Ala Leu Ser
        195                 200                 205

Tyr Asn Asn Ser Lys Gly Thr Ala Asn Trp Val Ala Trp Gln Leu Asn
    210                 215                 220

Ser Ser Trp Leu Gly Asn Ala Glu Arg Gln Asp Asn Phe Arg Pro Asp
225                 230                 235                 240

Lys Thr Leu Pro Ala Gly Trp Val Arg Val Thr Pro Ser Met Tyr Ser
                245                 250                 255

Gly Ser Gly Tyr Asp Arg Gly His Ile Ala Pro Ser Ala Asp Arg Thr
            260                 265                 270

Lys Thr Thr Glu Asp Asn Ala Ala Thr Phe Leu Met Thr Asn Met Met
        275                 280                 285

Pro Gln Thr Pro Asp Asn Asn Arg Asn Thr Trp Gly Asn Leu Glu Asp
    290                 295                 300

Tyr Cys Arg Glu Leu Val Ser Gln Gly Lys Glu Leu Tyr Ile Val Ala
305                 310                 315                 320

Gly Pro Asn Gly Ser Leu Gly Lys Pro Leu Lys Gly Lys Val Thr Val
                325                 330                 335
```

```
Pro Lys Ser Thr Trp Lys Ile Val Val Val Leu Asp Ser Pro Gly Ser
            340                 345                 350

Gly Leu Glu Gly Ile Thr Ala Asn Thr Arg Val Ile Ala Val Asn Ile
        355                 360                 365

Pro Asn Asp Pro Glu Leu Asn Asn Asp Trp Arg Ala Tyr Lys Val Ser
    370                 375                 380

Val Asp Glu Leu Glu Ser Leu Thr Gly Tyr Asp Phe Leu Ser Asn Val
385                 390                 395                 400

Ser Pro Asn Ile Gln Thr Ser Ile Glu Ser Lys Val Asp Asn Ala Ala
                405                 410                 415

Asp
```

<210> 434

<211> 1001
<212> PRT
<213> Artificial sequence

<220>
<223> cT11Avr_NucA

<400> 434

```
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1               5                   10                  15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
            20                  25                  30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
        35                  40                  45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
    50                  55                  60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65                  70                  75                  80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
                85                  90                  95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
            100                 105                 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val
    130                 135                 140
```

```
Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145                 150                 155                 160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
                165                 170                 175

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
            180                 185                 190

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        195                 200                 205

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
    210                 215                 220

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225                 230                 235                 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                245                 250                 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            260                 265                 270

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
        275                 280                 285

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
    290                 295                 300

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile
305                 310                 315                 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu
                325                 330                 335

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
            340                 345                 350

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
        355                 360                 365

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
    370                 375                 380

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
385                 390                 395                 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
                405                 410                 415
```

```
Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
            420                 425                 430

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
        435                 440                 445

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
    450                 455                 460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465                 470                 475                 480

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
                485                 490                 495

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
            500                 505                 510

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
        515                 520                 525

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
    530                 535                 540

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
545                 550                 555                 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
                565                 570                 575

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
            580                 585                 590

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
        595                 600                 605

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
    610                 615                 620

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
625                 630                 635                 640

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
                645                 650                 655

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
            660                 665                 670

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
        675                 680                 685
```

```
Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
    690                 695                 700

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
705                 710                 715                 720

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser
                725                 730                 735

Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Gly Ser Gln Val Pro Pro
            740                 745                 750

Leu Thr Glu Leu Ser Pro Ser Ile Ser Val His Leu Leu Leu Gly Asn
        755                 760                 765

Pro Ser Gly Ala Thr Pro Thr Lys Leu Thr Pro Asp Asn Tyr Leu Met
    770                 775                 780

Val Lys Asn Gln Tyr Ala Leu Ser Tyr Asn Asn Ser Lys Gly Thr Ala
785                 790                 795                 800

Asn Trp Val Ala Trp Gln Leu Asn Ser Ser Trp Leu Gly Asn Ala Glu
                805                 810                 815

Arg Gln Asp Asn Phe Arg Pro Asp Lys Thr Leu Pro Ala Gly Trp Val
            820                 825                 830

Arg Val Thr Pro Ser Met Tyr Ser Gly Ser Gly Tyr Asp Arg Gly His
        835                 840                 845

Ile Ala Pro Ser Ala Asp Arg Thr Lys Thr Thr Glu Asp Asn Ala Ala
    850                 855                 860

Thr Phe Leu Met Thr Asn Met Met Pro Gln Thr Pro Asp Asn Asn Arg
865                 870                 875                 880

Asn Thr Trp Gly Asn Leu Glu Asp Tyr Cys Arg Glu Leu Val Ser Gln
                885                 890                 895

Gly Lys Glu Leu Tyr Ile Val Ala Gly Pro Asn Gly Ser Leu Gly Lys
            900                 905                 910

Pro Leu Lys Gly Lys Val Thr Val Pro Lys Ser Thr Trp Lys Ile Val
        915                 920                 925

Val Val Leu Asp Ser Pro Gly Ser Gly Leu Glu Gly Ile Thr Ala Asn
    930                 935                 940

Thr Arg Val Ile Ala Val Asn Ile Pro Asn Asp Pro Glu Leu Asn Asn
945                 950                 955                 960
```

Asp Trp Arg Ala Tyr Lys Val Ser Val Asp Glu Leu Glu Ser Leu Thr
965 970 975

Gly Tyr Asp Phe Leu Ser Asn Val Ser Pro Asn Ile Gln Thr Ser Ile
980 985 990

Glu Ser Lys Val Asp Asn Ala Ala Asp
995 1000

<210> 435

<211> 303
<212> PRT
<213> Artificial sequence

<220>
<223> cT11_ColE7

<400> 435
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
100 105 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
115 120 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly
130 135 140

Ile Arg Ser Arg Ala Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro
145 150 155 160

Asp Pro Gly Ser Ser Gly Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr
165 170 175

Gly Lys Gly Lys Pro Val Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys
180 185 190

Asp Leu Gly Ser Pro Val Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp
195 200 205

Lys Glu Phe Lys Ser Phe Asp Asp Phe Arg Lys Lys Phe Trp Glu Glu
210 215 220

Val Ser Lys Asp Pro Glu Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn
225 230 235 240

Asp Arg Met Lys Val Gly Lys Ala Pro Lys Thr Arg Thr Gln Asp Val
245 250 255

Ser Gly Lys Arg Thr Ser Phe Glu Leu His His Glu Lys Pro Ile Ser
260 265 270

Gln Asn Gly Gly Val Tyr Asp Met Asp Asn Ile Ser Val Val Thr Pro
275 280 285

Lys Arg His Ile Asp Ile His Arg Gly Lys Gly Ser Ser Ala Asp
290 295 300

<210> 436

<211> 887
<212> PRT
<213> Artificial sequence

<220>
<223> cT11Avr_ColE7

<400> 436
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
100 105 110

```
Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val
    130                 135                 140

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145                 150                 155                 160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
                165                 170                 175

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
            180                 185                 190

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        195                 200                 205

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
    210                 215                 220

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225                 230                 235                 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                245                 250                 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            260                 265                 270

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
        275                 280                 285

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
    290                 295                 300

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile
305                 310                 315                 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu
                325                 330                 335

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
            340                 345                 350

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
        355                 360                 365

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
    370                 375                 380
```

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
385 390 395 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
405 410 415

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
420 425 430

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
435 440 445

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
450 455 460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465 470 475 480

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
485 490 495

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
500 505 510

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
515 520 525

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
530 535 540

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
545 550 555 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
565 570 575

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
580 585 590

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
595 600 605

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
610 615 620

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
625 630 635 640

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
645 650 655

```
Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
            660                 665                 670

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
        675                 680                 685

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
    690                 695                 700

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
705                 710                 715                 720

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser
                725                 730                 735

Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Gly Ser Ser Gly Ser Lys
            740                 745                 750

Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys Gly Lys Pro Val Asn Asn
        755                 760                 765

Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu Gly Ser Pro Val Pro Asp
    770                 775                 780

Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu Phe Lys Ser Phe Asp Asp
785                 790                 795                 800

Phe Arg Lys Lys Phe Trp Glu Glu Val Ser Lys Asp Pro Glu Leu Ser
                805                 810                 815

Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg Met Lys Val Gly Lys Ala
            820                 825                 830

Pro Lys Thr Arg Thr Gln Asp Val Ser Gly Lys Arg Thr Ser Phe Glu
        835                 840                 845

Leu His His Glu Lys Pro Ile Ser Gln Asn Gly Gly Val Tyr Asp Met
    850                 855                 860

Asp Asn Ile Ser Val Val Thr Pro Lys Arg His Ile Asp Ile His Arg
865                 870                 875                 880

Gly Lys Gly Ser Ser Ala Asp
                885
```

<210> 437

<211> 819
<212> PRT
<213> Artificial sequence

<220>
<223> cT11NptIIT5-L_ColE7_A497_CFS1

<400> 437

```
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1               5                  10                  15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
            20                  25                  30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
        35                  40                  45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
    50                  55                  60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65                  70                  75                  80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
                85                  90                  95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
            100                 105                 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val
    130                 135                 140

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145                 150                 155                 160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                165                 170                 175

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
            180                 185                 190

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        195                 200                 205

Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu
    210                 215                 220

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225                 230                 235                 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                245                 250                 255
```

```
Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            260                 265                 270

Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly
        275                 280                 285

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    290                 295                 300

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
305                 310                 315                 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
                325                 330                 335

Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser
            340                 345                 350

Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
        355                 360                 365

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
    370                 375                 380

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
385                 390                 395                 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
                405                 410                 415

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
            420                 425                 430

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
        435                 440                 445

Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr
    450                 455                 460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465                 470                 475                 480

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
                485                 490                 495

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
            500                 505                 510

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
        515                 520                 525
```

```
Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
    530                 535                 540

Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly
545                 550                 555                 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
                565                 570                 575

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
            580                 585                 590

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
        595                 600                 605

Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser
    610                 615                 620

Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
625                 630                 635                 640

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
                645                 650                 655

Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln
            660                 665                 670

Leu Ser Arg Pro Asp Pro Gly Ser Ser Gly Ser Lys Arg Asn Lys Pro
        675                 680                 685

Gly Lys Ala Thr Gly Lys Gly Lys Pro Val Asn Asn Lys Trp Leu Asn
    690                 695                 700

Asn Ala Gly Lys Asp Leu Gly Ser Pro Val Pro Asp Arg Ile Ala Asn
705                 710                 715                 720

Lys Leu Arg Asp Lys Glu Phe Lys Ser Phe Asp Asp Phe Arg Ala Lys
                725                 730                 735

Phe Trp Glu Glu Val Ser Lys Asp Pro Glu Leu Ser Lys Gln Phe Ser
            740                 745                 750

Arg Asn Asn Asn Asp Arg Met Lys Val Gly Lys Ala Pro Lys Thr Arg
        755                 760                 765

Thr Gln Asp Val Ser Gly Lys Arg Thr Ser Phe Glu Leu His His Glu
    770                 775                 780

Lys Pro Ile Ser Gln Asn Gly Gly Val Tyr Asp Met Asp Asn Ile Ser
785                 790                 795                 800
```

Val Val Thr Pro Lys Arg His Ile Asp Ile His Arg Gly Lys Gly Ser
805 810 815

Thr Gly Asp

<210> 438

<211> 819
<212> PRT
<213> Artificial sequence

<220>
<223> cT11NptIIT6-L_ColE7_A497_CFS1

<400> 438
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
100 105 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
115 120 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Gln Gln Val Val
130 135 140

Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145 150 155 160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
165 170 175

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
180 185 190

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
195 200 205

```
Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
    210                 215                 220

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225                 230                 235                 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln
                245                 250                 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            260                 265                 270

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
        275                 280                 285

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    290                 295                 300

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn
305                 310                 315                 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
                325                 330                 335

Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser
            340                 345                 350

Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
        355                 360                 365

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
    370                 375                 380

Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
385                 390                 395                 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
                405                 410                 415

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
            420                 425                 430

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
        435                 440                 445

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
    450                 455                 460

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465                 470                 475                 480
```

```
Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu
                485                 490                 495

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
            500                 505                 510

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln
        515                 520                 525

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
    530                 535                 540

Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly
545                 550                 555                 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
                565                 570                 575

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile
            580                 585                 590

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu
        595                 600                 605

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
    610                 615                 620

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
625                 630                 635                 640

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
                645                 650                 655

Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln
            660                 665                 670

Leu Ser Arg Pro Asp Pro Gly Ser Ser Gly Ser Lys Arg Asn Lys Pro
        675                 680                 685

Gly Lys Ala Thr Gly Lys Gly Lys Pro Val Asn Asn Lys Trp Leu Asn
    690                 695                 700

Asn Ala Gly Lys Asp Leu Gly Ser Pro Val Pro Asp Arg Ile Ala Asn
705                 710                 715                 720

Lys Leu Arg Asp Lys Glu Phe Lys Ser Phe Asp Asp Phe Arg Ala Lys
                725                 730                 735

Phe Trp Glu Glu Val Ser Lys Asp Pro Glu Leu Ser Lys Gln Phe Ser
            740                 745                 750
```

```
Arg Asn Asn Asn Asp Arg Met Lys Val Gly Lys Ala Pro Lys Thr Arg
        755                 760                 765

Thr Gln Asp Val Ser Gly Lys Arg Thr Ser Phe Glu Leu His His Glu
    770                 775                 780

Lys Pro Ile Ser Gln Asn Gly Gly Val Tyr Asp Met Asp Asn Ile Ser
785                 790                 795                 800

Val Val Thr Pro Lys Arg His Ile Asp Ile His Arg Gly Lys Gly Ser
                805                 810                 815

Thr Gly Asp
```

<210> 439

<211> 529
<212> PRT
<213> Artificial sequence

<220>
<223> cT11_scTB2aD01

<400> 439

```
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1               5                   10                  15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
            20                  25                  30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
        35                  40                  45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
    50                  55                  60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65                  70                  75                  80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
                85                  90                  95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
            100                 105                 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly
    130                 135                 140
```

```
Ile Arg Ser Arg Ala Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro
145                 150                 155                 160

Asp Pro Gly Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu Lys Met Lys
                165                 170                 175

Gly Gln Gly Pro Ser Gly Asn Thr Lys Tyr Asn Glu Glu Phe Leu Leu
            180                 185                 190

Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile Val Ala Gln Ile
        195                 200                 205

Lys Pro His Gln Ser Cys Lys Phe Lys His Gln Leu Arg Leu Thr Phe
    210                 215                 220

Asn Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val
225                 230                 235                 240

Asp Glu Ile Gly Val Gly His Val Tyr Asp Ser Gly Ser Val Ser Tyr
                245                 250                 255

Tyr Gln Leu Ser Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu
            260                 265                 270

Gln Pro Phe Leu Glu Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys
        275                 280                 285

Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu
    290                 295                 300

Glu Val Cys Thr Trp Val Asp Gln Val Ala Ala Leu Asn Asp Ser Lys
305                 310                 315                 320

Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu
                325                 330                 335

Ser Glu Lys Lys Lys Ser Ser Pro Ala Ala Gly Asp Ser Ser Val Ser
            340                 345                 350

Asn Ser Glu His Ile Ala Pro Leu Ser Leu Pro Ser Ser Pro Pro Ser
        355                 360                 365

Val Gly Ser Asn Lys Lys Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp
    370                 375                 380

Ser Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys
385                 390                 395                 400

Phe Lys His Tyr Leu Ser Leu Thr Phe Cys Val Thr Gln Lys Thr Gln
                405                 410                 415
```

Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Arg Ile Gly Val Gly Tyr
420 425 430

Val Gln Asp Arg Gly Ser Val Ser Asp Tyr Arg Leu Ser Glu Ile Lys
435 440 445

Pro Leu Arg Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys
450 455 460

Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser
465 470 475 480

Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp
485 490 495

Gln Val Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu
500 505 510

Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser
515 520 525

Pro

<210> 440

<211> 529
<212> PRT
<213> Artificial sequence

<220>
<223> cT11_scTB2aD01_148C

<400> 440
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
100 105 110

```
Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly
    130                 135                 140

Ile Arg Ser Arg Ala Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro
145                 150                 155                 160

Asp Pro Gly Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu Lys Met Lys
                165                 170                 175

Gly Gln Gly Pro Ser Gly Asn Thr Lys Tyr Asn Glu Glu Phe Leu Leu
            180                 185                 190

Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile Val Ala Gln Ile
        195                 200                 205

Lys Pro His Gln Ser Cys Lys Phe Lys His Gln Leu Arg Leu Thr Phe
    210                 215                 220

Asn Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val
225                 230                 235                 240

Asp Glu Ile Gly Val Gly His Val Tyr Asp Ser Gly Ser Val Ser Tyr
                245                 250                 255

Tyr Gln Leu Ser Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu
            260                 265                 270

Gln Pro Phe Leu Glu Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys
        275                 280                 285

Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu
    290                 295                 300

Glu Val Cys Thr Trp Val Asp Gln Val Ala Ala Leu Asn Asp Ser Lys
305                 310                 315                 320

Thr Arg Lys Thr Thr Ser Glu Thr Cys Arg Ala Val Leu Asp Ser Leu
                325                 330                 335

Ser Glu Lys Lys Lys Ser Ser Pro Ala Ala Gly Asp Ser Ser Val Ser
            340                 345                 350

Asn Ser Glu His Ile Ala Pro Leu Ser Leu Pro Ser Ser Pro Pro Ser
        355                 360                 365

Val Gly Ser Asn Lys Lys Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp
    370                 375                 380
```

Ser Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys
385 390 395 400

Phe Lys His Tyr Leu Ser Leu Thr Phe Cys Val Thr Gln Lys Thr Gln
405 410 415

Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Arg Ile Gly Val Gly Tyr
420 425 430

Val Gln Asp Arg Gly Ser Val Ser Asp Tyr Arg Leu Ser Glu Ile Lys
435 440 445

Pro Leu Arg Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys
450 455 460

Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser
465 470 475 480

Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp
485 490 495

Gln Val Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu
500 505 510

Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser
515 520 525

Pro

<210> 441

<211> 529
<212> PRT
<213> Artificial sequence

<220>
<223> cT11_scTB2aD01_333C

<400> 441
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

```
Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65                  70                  75                  80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
                85                  90                  95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
            100                 105                 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly
    130                 135                 140

Ile Arg Ser Arg Ala Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro
145                 150                 155                 160

Asp Pro Gly Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu Lys Met Lys
                165                 170                 175

Gly Gln Gly Pro Ser Gly Asn Thr Lys Tyr Asn Glu Glu Phe Leu Leu
            180                 185                 190

Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile Val Ala Gln Ile
        195                 200                 205

Lys Pro His Gln Ser Cys Lys Phe Lys His Gln Leu Arg Leu Thr Phe
    210                 215                 220

Asn Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val
225                 230                 235                 240

Asp Glu Ile Gly Val Gly His Val Tyr Asp Ser Gly Ser Val Ser Tyr
                245                 250                 255

Tyr Gln Leu Ser Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu
            260                 265                 270

Gln Pro Phe Leu Glu Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys
        275                 280                 285

Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu
    290                 295                 300

Glu Val Cys Thr Trp Val Asp Gln Val Ala Ala Leu Asn Asp Ser Lys
305                 310                 315                 320

Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu
                325                 330                 335
```

Ser Glu Lys Lys Lys Ser Ser Pro Ala Ala Gly Asp Ser Ser Val Ser
340 345 350

Asn Ser Glu His Ile Ala Pro Leu Ser Leu Pro Ser Ser Pro Pro Ser
355 360 365

Val Gly Ser Asn Lys Lys Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp
370 375 380

Ser Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys
385 390 395 400

Phe Lys His Tyr Leu Ser Leu Thr Phe Cys Val Thr Gln Lys Thr Gln
405 410 415

Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Arg Ile Gly Val Gly Tyr
420 425 430

Val Gln Asp Arg Gly Ser Val Ser Asp Tyr Arg Leu Ser Glu Ile Lys
435 440 445

Pro Leu Arg Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys
450 455 460

Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser
465 470 475 480

Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp
485 490 495

Gln Val Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu
500 505 510

Thr Cys Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser
515 520 525

Pro

<210> 442

<211> 349
<212> PRT
<213> Artificial sequence

<220>
<223> TCRB02-A-V148C

<400> 442
Met Ala Asn Thr Lys Tyr Asn Glu Glu Phe Leu Leu Tyr Leu Ala Gly
1 5 10 15

Phe Val Asp Gly Asp Gly Ser Ile Val Ala Gln Ile Lys Pro His Gln
20 25 30

```
Ser Cys Lys Phe Lys His Gln Leu Arg Leu Thr Phe Asn Val Thr Gln
        35                  40                  45

Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
    50                  55                  60

Val Gly His Val Tyr Asp Ser Gly Ser Val Ser Tyr Tyr Gln Leu Ser
65                  70                  75                  80

Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu
                85                  90                  95

Glu Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln
            100                 105                 110

Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr
        115                 120                 125

Trp Val Asp Gln Val Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr
    130                 135                 140

Thr Ser Glu Thr Cys Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys
145                 150                 155                 160

Lys Ser Ser Pro Ala Ala Gly Asp Ser Ser Val Ser Asn Ser Glu His
                165                 170                 175

Ile Ala Pro Leu Ser Leu Pro Ser Ser Pro Pro Ser Val Gly Ser Asn
            180                 185                 190

Lys Lys Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Ser Asp Gly Ser
        195                 200                 205

Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Tyr
    210                 215                 220

Leu Ser Leu Thr Phe Cys Val Thr Gln Lys Thr Gln Arg Arg Trp Phe
225                 230                 235                 240

Leu Asp Lys Leu Val Asp Arg Ile Gly Val Gly Tyr Val Gln Asp Arg
                245                 250                 255

Gly Ser Val Ser Asp Tyr Arg Leu Ser Glu Ile Lys Pro Leu Arg Asn
            260                 265                 270

Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala
        275                 280                 285

Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser
    290                 295                 300
```

Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Val Ala Ala
305 310 315 320

Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala
325 330 335

Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro
340 345

<210> 443

<211> 349
<212> PRT
<213> Artificial sequence

<220>
<223> TCRB02-A-V333C

<400> 443
Met Ala Asn Thr Lys Tyr Asn Glu Glu Phe Leu Leu Tyr Leu Ala Gly
1 5 10 15

Phe Val Asp Gly Asp Gly Ser Ile Val Ala Gln Ile Lys Pro His Gln
20 25 30

Ser Cys Lys Phe Lys His Gln Leu Arg Leu Thr Phe Asn Val Thr Gln
35 40 45

Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
50 55 60

Val Gly His Val Tyr Asp Ser Gly Ser Val Ser Tyr Tyr Gln Leu Ser
65 70 75 80

Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu
85 90 95

Glu Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln
100 105 110

Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr
115 120 125

Trp Val Asp Gln Val Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr
130 135 140

Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys
145 150 155 160

Lys Ser Ser Pro Ala Ala Gly Asp Ser Ser Val Ser Asn Ser Glu His
165 170 175

Ile Ala Pro Leu Ser Leu Pro Ser Ser Pro Pro Ser Val Gly Ser Asn
180 185 190

Lys Lys Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Ser Asp Gly Ser
195 200 205

Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Tyr
210 215 220

Leu Ser Leu Thr Phe Cys Val Thr Gln Lys Thr Gln Arg Arg Trp Phe
225 230 235 240

Leu Asp Lys Leu Val Asp Arg Ile Gly Val Gly Tyr Val Gln Asp Arg
245 250 255

Gly Ser Val Ser Asp Tyr Arg Leu Ser Glu Ile Lys Pro Leu Arg Asn
260 265 270

Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala
275 280 285

Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser
290 295 300

Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Val Ala Ala
305 310 315 320

Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Cys Arg Ala
325 330 335

Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro
340 345

<210> 444

<211> 977
<212> PRT
<213> Artificial sequence

<220>
<223> cT11TB2A1_scTB2aD01

<400> 444
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

```
Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65                  70                  75                  80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
                85                  90                  95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
            100                 105                 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Gln Gln Val Val
    130                 135                 140

Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145                 150                 155                 160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                165                 170                 175

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
            180                 185                 190

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        195                 200                 205

Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu
    210                 215                 220

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225                 230                 235                 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln
                245                 250                 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            260                 265                 270

Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly
        275                 280                 285

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    290                 295                 300

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
305                 310                 315                 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
                325                 330                 335
```

```
Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
            340                 345                 350

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
        355                 360                 365

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
    370                 375                 380

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
385                 390                 395                 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val
                405                 410                 415

Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
            420                 425                 430

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
        435                 440                 445

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
    450                 455                 460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465                 470                 475                 480

Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu
                485                 490                 495

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
            500                 505                 510

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
        515                 520                 525

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
    530                 535                 540

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
545                 550                 555                 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
                565                 570                 575

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
            580                 585                 590

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
        595                 600                 605
```

```
Asp Pro Gly Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu Lys Met Lys
    610                 615                 620

Gly Gln Gly Pro Ser Gly Asn Thr Lys Tyr Asn Glu Glu Phe Leu Leu
625                 630                 635                 640

Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile Val Ala Gln Ile
                645                 650                 655

Lys Pro His Gln Ser Cys Lys Phe Lys His Gln Leu Arg Leu Thr Phe
            660                 665                 670

Asn Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val
        675                 680                 685

Asp Glu Ile Gly Val Gly His Val Tyr Asp Ser Gly Ser Val Ser Tyr
    690                 695                 700

Tyr Gln Leu Ser Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu
705                 710                 715                 720

Gln Pro Phe Leu Glu Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys
                725                 730                 735

Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu
            740                 745                 750

Glu Val Cys Thr Trp Val Asp Gln Val Ala Ala Leu Asn Asp Ser Lys
        755                 760                 765

Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu
    770                 775                 780

Ser Glu Lys Lys Lys Ser Ser Pro Ala Ala Gly Asp Ser Ser Val Ser
785                 790                 795                 800

Asn Ser Glu His Ile Ala Pro Leu Ser Leu Pro Ser Ser Pro Pro Ser
                805                 810                 815

Val Gly Ser Asn Lys Lys Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp
            820                 825                 830

Ser Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys
        835                 840                 845

Phe Lys His Tyr Leu Ser Leu Thr Phe Cys Val Thr Gln Lys Thr Gln
    850                 855                 860

Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Arg Ile Gly Val Gly Tyr
865                 870                 875                 880
```

Val Gln Asp Arg Gly Ser Val Ser Asp Tyr Arg Leu Ser Glu Ile Lys
885 890 895

Pro Leu Arg Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys
900 905 910

Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser
915 920 925

Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp
930 935 940

Gln Val Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu
945 950 955 960

Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser
965 970 975

Pro

<210> 445

<211> 977
<212> PRT
<213> Artificial sequence

<220>
<223> cT11TB2A2_scTB2aD01

<400> 445
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
100 105 110

```
Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Gln Gln Val Val
    130                 135                 140

Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145                 150                 155                 160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                165                 170                 175

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
            180                 185                 190

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        195                 200                 205

Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu
    210                 215                 220

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225                 230                 235                 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln
                245                 250                 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            260                 265                 270

Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly
        275                 280                 285

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    290                 295                 300

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn
305                 310                 315                 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
                325                 330                 335

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
            340                 345                 350

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
        355                 360                 365

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
    370                 375                 380
```

```
Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
385                 390                 395                 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val
                405                 410                 415

Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
            420                 425                 430

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
        435                 440                 445

Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr
    450                 455                 460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465                 470                 475                 480

Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu
                485                 490                 495

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
            500                 505                 510

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
        515                 520                 525

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
    530                 535                 540

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
545                 550                 555                 560

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
                565                 570                 575

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
            580                 585                 590

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
        595                 600                 605

Asp Pro Gly Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu Lys Met Lys
    610                 615                 620

Gly Gln Gly Pro Ser Gly Asn Thr Lys Tyr Asn Glu Glu Phe Leu Leu
625                 630                 635                 640

Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile Val Ala Gln Ile
                645                 650                 655
```

```
Lys Pro His Gln Ser Cys Lys Phe Lys His Gln Leu Arg Leu Thr Phe
            660                 665                 670

Asn Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val
        675                 680                 685

Asp Glu Ile Gly Val Gly His Val Tyr Asp Ser Gly Ser Val Ser Tyr
    690                 695                 700

Tyr Gln Leu Ser Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu
705                 710                 715                 720

Gln Pro Phe Leu Glu Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys
                725                 730                 735

Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu
            740                 745                 750

Glu Val Cys Thr Trp Val Asp Gln Val Ala Ala Leu Asn Asp Ser Lys
        755                 760                 765

Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu
    770                 775                 780

Ser Glu Lys Lys Lys Ser Ser Pro Ala Ala Gly Asp Ser Ser Val Ser
785                 790                 795                 800

Asn Ser Glu His Ile Ala Pro Leu Ser Leu Pro Ser Ser Pro Pro Ser
                805                 810                 815

Val Gly Ser Asn Lys Lys Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp
            820                 825                 830

Ser Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys
        835                 840                 845

Phe Lys His Tyr Leu Ser Leu Thr Phe Cys Val Thr Gln Lys Thr Gln
    850                 855                 860

Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Arg Ile Gly Val Gly Tyr
865                 870                 875                 880

Val Gln Asp Arg Gly Ser Val Ser Asp Tyr Arg Leu Ser Glu Ile Lys
                885                 890                 895

Pro Leu Arg Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys
            900                 905                 910

Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser
        915                 920                 925
```

Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp
    930   935   940

Gln Val Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu
945   950   955   960

Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser
                965   970   975

Pro

<210> 446

<211> 977
<212> PRT
<213> Artificial sequence

<220>
<223> cT11TB2A3_scTB2aD01

<400> 446
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1   5   10   15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
            20   25   30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
        35   40   45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
    50   55   60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65   70   75   80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
                85   90   95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
            100   105   110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115   120   125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Gln Gln Val Val
    130   135   140

Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145   150   155   160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
                165   170   175

```
Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
            180                 185                 190

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        195                 200                 205

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
    210                 215                 220

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225                 230                 235                 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                245                 250                 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            260                 265                 270

Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly
        275                 280                 285

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    290                 295                 300

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
305                 310                 315                 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
                325                 330                 335

Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser
            340                 345                 350

Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
        355                 360                 365

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
    370                 375                 380

Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
385                 390                 395                 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val
                405                 410                 415

Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
            420                 425                 430

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
        435                 440                 445
```

```
Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr
    450                 455                 460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465                 470                 475                 480

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
                485                 490                 495

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
            500                 505                 510

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
        515                 520                 525

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
    530                 535                 540

Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly
545                 550                 555                 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
                565                 570                 575

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
            580                 585                 590

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
        595                 600                 605

Asp Pro Gly Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu Lys Met Lys
    610                 615                 620

Gly Gln Gly Pro Ser Gly Asn Thr Lys Tyr Asn Glu Glu Phe Leu Leu
625                 630                 635                 640

Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile Val Ala Gln Ile
                645                 650                 655

Lys Pro His Gln Ser Cys Lys Phe Lys His Gln Leu Arg Leu Thr Phe
            660                 665                 670

Asn Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val
        675                 680                 685

Asp Glu Ile Gly Val Gly His Val Tyr Asp Ser Gly Ser Val Ser Tyr
    690                 695                 700

Tyr Gln Leu Ser Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu
705                 710                 715                 720
```

```
Gln Pro Phe Leu Glu Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys
                725                 730                 735

Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu
            740                 745                 750

Glu Val Cys Thr Trp Val Asp Gln Val Ala Ala Leu Asn Asp Ser Lys
        755                 760                 765

Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu
    770                 775                 780

Ser Glu Lys Lys Lys Ser Ser Pro Ala Ala Gly Asp Ser Ser Val Ser
785                 790                 795                 800

Asn Ser Glu His Ile Ala Pro Leu Ser Leu Pro Ser Ser Pro Pro Ser
                805                 810                 815

Val Gly Ser Asn Lys Lys Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp
            820                 825                 830

Ser Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys
        835                 840                 845

Phe Lys His Tyr Leu Ser Leu Thr Phe Cys Val Thr Gln Lys Thr Gln
    850                 855                 860

Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Arg Ile Gly Val Gly Tyr
865                 870                 875                 880

Val Gln Asp Arg Gly Ser Val Ser Asp Tyr Arg Leu Ser Glu Ile Lys
                885                 890                 895

Pro Leu Arg Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys
            900                 905                 910

Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser
        915                 920                 925

Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp
    930                 935                 940

Gln Val Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu
945                 950                 955                 960

Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser
                965                 970                 975

Pro
```

<210> 447

<211> 1146
<212> PRT
<213> Artificial sequence

<220>
<223> TevD02_cT11Avr_FokI-L

<400> 447
```
Met Ala Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys
1               5                   10                  15


Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His
            20                  25                  30


Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg
        35                  40                  45


Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu
    50                  55                  60


Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile
65                  70                  75                  80


Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr
                85                  90                  95


Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys
            100                 105                 110


Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp
        115                 120                 125


Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp
    130                 135                 140


Asn Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr
145                 150                 155                 160


Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn
                165                 170                 175


Ser Phe Phe Asn His Lys His Ser Gln Gly Pro Ser Gly Val Asp Leu
            180                 185                 190


Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys
        195                 200                 205


Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly
    210                 215                 220
```

```
Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu
225                 230                 235                 240

Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu
                245                 250                 255

Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala
            260                 265                 270

Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro
        275                 280                 285

Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly
    290                 295                 300

Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr
305                 310                 315                 320

Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
                325                 330                 335

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
            340                 345                 350

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
        355                 360                 365

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
    370                 375                 380

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
385                 390                 395                 400

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
                405                 410                 415

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
            420                 425                 430

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
        435                 440                 445

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
    450                 455                 460

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
465                 470                 475                 480

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
                485                 490                 495
```

```
Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
            500                 505                 510

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
        515                 520                 525

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
    530                 535                 540

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
545                 550                 555                 560

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
                565                 570                 575

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            580                 585                 590

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        595                 600                 605

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
    610                 615                 620

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
625                 630                 635                 640

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                645                 650                 655

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            660                 665                 670

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
        675                 680                 685

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
    690                 695                 700

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
705                 710                 715                 720

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                725                 730                 735

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
            740                 745                 750

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        755                 760                 765
```

```
Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    770                 775                 780

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
785                 790                 795                 800

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
                805                 810                 815

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            820                 825                 830

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
        835                 840                 845

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
    850                 855                 860

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
865                 870                 875                 880

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
                885                 890                 895

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
            900                 905                 910

Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln
        915                 920                 925

Leu Ser Arg Pro Asp Pro Gly Ser Ser Gly Pro Asn Arg Gly Val Thr
    930                 935                 940

Lys Gln Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg
945                 950                 955                 960

His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile
                965                 970                 975

Ala Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu
            980                 985                 990

Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser
        995                 1000                1005

Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr
    1010                1015                1020

Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro
1025                1030                1035                1040
```

Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn Gln Thr
1045 1050 1055

Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser
1060 1065 1070

Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly
1075 1080 1085

Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn
1090 1095 1100

Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile
1105 1110 1115 1120

Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn
1125 1130 1135

Gly Glu Ile Asn Phe Gly Ser Ser Ala Asp
1140 1145

<210> 448

<211> 1124
<212> PRT
<213> Artificial sequence

<220>
<223> TevD02_cT11Avr_TevD02

<400> 448
Met Ala Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys
1 5 10 15

Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His
20 25 30

Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg
35 40 45

Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu
50 55 60

Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile
65 70 75 80

Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr
85 90 95

Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys
100 105 110

Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp
115 120 125

```
Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp
    130                 135                 140

Asn Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr
145                 150                 155                 160

Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn
                165                 170                 175

Ser Phe Phe Asn His Lys His Ser Gln Gly Pro Ser Gly Val Asp Leu
            180                 185                 190

Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys
        195                 200                 205

Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly
    210                 215                 220

Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu
225                 230                 235                 240

Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu
                245                 250                 255

Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala
            260                 265                 270

Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro
        275                 280                 285

Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly
    290                 295                 300

Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr
305                 310                 315                 320

Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
                325                 330                 335

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
            340                 345                 350

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
        355                 360                 365

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
    370                 375                 380

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
385                 390                 395                 400
```

```
Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
                405                 410                 415

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
            420                 425                 430

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
        435                 440                 445

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
    450                 455                 460

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
465                 470                 475                 480

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
                485                 490                 495

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
            500                 505                 510

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
        515                 520                 525

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
    530                 535                 540

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
545                 550                 555                 560

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
                565                 570                 575

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            580                 585                 590

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        595                 600                 605

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
    610                 615                 620

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
625                 630                 635                 640

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                645                 650                 655

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            660                 665                 670
```

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
675 680 685

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
690 695 700

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
705 710 715 720

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
725 730 735

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
740 745 750

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
755 760 765

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
770 775 780

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
785 790 795 800

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
805 810 815

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
820 825 830

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
835 840 845

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
850 855 860

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
865 870 875 880

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
885 890 895

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
900 905 910

Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln
915 920 925

Leu Ser Arg Pro Asp Pro Gly Ser Lys Ser Gly Ile Tyr Gln Ile Lys
930 935 940

Asn Thr Leu Asn Asn Lys Val Tyr Val Gly Ser Ala Lys Asp Phe Glu
945 950 955 960

Lys Arg Trp Lys Arg His Phe Lys Asp Leu Glu Lys Gly Cys His Ser
965 970 975

Ser Ile Lys Leu Gln Arg Ser Phe Asn Lys His Gly Asn Val Phe Glu
980 985 990

Cys Ser Ile Leu Glu Glu Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu
995 1000 1005

Arg Glu Asn Phe Trp Ile Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr
1010 1015 1020

Asn Ile Ala Asp Ala Thr Phe Gly Asp Thr Cys Ser Thr His Pro Leu
1025 1030 1035 1040

Lys Glu Glu Ile Ile Lys Lys Arg Ser Glu Thr Val Lys Ala Lys Met
1045 1050 1055

Leu Lys Leu Gly Pro Asp Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly
1060 1065 1070

Ser Lys Asn Gly Arg Trp Asn Pro Glu Thr His Lys Phe Cys Lys Cys
1075 1080 1085

Gly Val Arg Ile Gln Thr Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn
1090 1095 1100

Arg Ser Gly Glu Asn Asn Ser Phe Phe Asn His Lys His Ser Gln Gly
1105 1110 1115 1120

Pro Ser Ala Asp

<210> 449

<211> 1146
<212> PRT
<213> Artificial sequence

<220>
<223> TevD02_cT11Avr_FokI-L

<400> 449
Met Ala Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys
1 5 10 15

Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His
20 25 30

```
Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg
        35                  40                  45

Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu
    50                  55                  60

Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile
65                  70                  75                  80

Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr
                85                  90                  95

Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys
            100                 105                 110

Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp
        115                 120                 125

Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp
    130                 135                 140

Asn Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr
145                 150                 155                 160

Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn
                165                 170                 175

Ser Phe Phe Asn His Lys His Ser Gln Gly Pro Ser Gly Val Asp Leu
            180                 185                 190

Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys
        195                 200                 205

Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly
    210                 215                 220

Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu
225                 230                 235                 240

Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu
                245                 250                 255

Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala
            260                 265                 270

Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro
        275                 280                 285

Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly
    290                 295                 300
```

```
Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr
305                 310                 315                 320

Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
                325                 330                 335

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
            340                 345                 350

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
        355                 360                 365

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
    370                 375                 380

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
385                 390                 395                 400

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
                405                 410                 415

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
            420                 425                 430

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
        435                 440                 445

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
    450                 455                 460

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
465                 470                 475                 480

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
                485                 490                 495

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
            500                 505                 510

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
        515                 520                 525

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
    530                 535                 540

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
545                 550                 555                 560

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
                565                 570                 575
```

```
Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            580                 585                 590

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        595                 600                 605

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
    610                 615                 620

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
625                 630                 635                 640

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                645                 650                 655

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            660                 665                 670

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
        675                 680                 685

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
    690                 695                 700

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
705                 710                 715                 720

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                725                 730                 735

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
            740                 745                 750

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        755                 760                 765

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    770                 775                 780

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
785                 790                 795                 800

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
                805                 810                 815

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            820                 825                 830

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
        835                 840                 845
```

```
Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
    850                 855                 860

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
865                 870                 875                 880

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
                885                 890                 895

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
            900                 905                 910

Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln
        915                 920                 925

Leu Ser Arg Pro Asp Pro Gly Ser Ser Gly Pro Asn Arg Gly Val Thr
    930                 935                 940

Lys Gln Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu Leu Arg
945                 950                 955                 960

His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile
                965                 970                 975

Ala Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu
            980                 985                 990

Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser
        995                 1000                1005

Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr
    1010                1015                1020

Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro
1025                1030                1035                1040

Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn Gln Thr
                1045                1050                1055

Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser
            1060                1065                1070

Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly
        1075                1080                1085

Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn
    1090                1095                1100

Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile
1105                1110                1115                1120
```

Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn
1125 1130 1135

Gly Glu Ile Asn Phe Gly Ser Ser Ala Asp
1140 1145

<210> 450

<211> 1124
<212> PRT
<213> Artificial sequence

<220>
<223> TevD02_cT11Avr_TevD02

<400> 450
Met Ala Lys Ser Gly Ile Tyr Gln Ile Lys Asn Thr Leu Asn Asn Lys
1 5 10 15

Val Tyr Val Gly Ser Ala Lys Asp Phe Glu Lys Arg Trp Lys Arg His
20 25 30

Phe Lys Asp Leu Glu Lys Gly Cys His Ser Ser Ile Lys Leu Gln Arg
35 40 45

Ser Phe Asn Lys His Gly Asn Val Phe Glu Cys Ser Ile Leu Glu Glu
50 55 60

Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu Arg Glu Asn Phe Trp Ile
65 70 75 80

Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr Asn Ile Ala Asp Ala Thr
85 90 95

Phe Gly Asp Thr Cys Ser Thr His Pro Leu Lys Glu Glu Ile Ile Lys
100 105 110

Lys Arg Ser Glu Thr Val Lys Ala Lys Met Leu Lys Leu Gly Pro Asp
115 120 125

Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly Ser Lys Asn Gly Arg Trp
130 135 140

Asn Pro Glu Thr His Lys Phe Cys Lys Cys Gly Val Arg Ile Gln Thr
145 150 155 160

Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn Arg Ser Gly Glu Asn Asn
165 170 175

Ser Phe Phe Asn His Lys His Ser Gln Gly Pro Ser Gly Val Asp Leu
180 185 190

Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys
195 200 205

```
Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly
    210                 215                 220

Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu
225                 230                 235                 240

Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu
                245                 250                 255

Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala
            260                 265                 270

Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro
        275                 280                 285

Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly
    290                 295                 300

Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr
305                 310                 315                 320

Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
                325                 330                 335

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
            340                 345                 350

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
        355                 360                 365

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
    370                 375                 380

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
385                 390                 395                 400

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
                405                 410                 415

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
            420                 425                 430

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
        435                 440                 445

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
    450                 455                 460

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
465                 470                 475                 480
```

```
Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
                485                 490                 495

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
            500                 505                 510

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
        515                 520                 525

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
    530                 535                 540

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
545                 550                 555                 560

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
                565                 570                 575

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
            580                 585                 590

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
        595                 600                 605

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
    610                 615                 620

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
625                 630                 635                 640

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
                645                 650                 655

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
            660                 665                 670

Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
        675                 680                 685

Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
    690                 695                 700

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
705                 710                 715                 720

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
                725                 730                 735

Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu
            740                 745                 750
```

```
Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
        755                 760                 765

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
    770                 775                 780

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
785                 790                 795                 800

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
                805                 810                 815

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
            820                 825                 830

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
        835                 840                 845

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
    850                 855                 860

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
865                 870                 875                 880

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
                885                 890                 895

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
            900                 905                 910

Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln
        915                 920                 925

Leu Ser Arg Pro Asp Pro Gly Ser Lys Ser Gly Ile Tyr Gln Ile Lys
    930                 935                 940

Asn Thr Leu Asn Asn Lys Val Tyr Val Gly Ser Ala Lys Asp Phe Glu
945                 950                 955                 960

Lys Arg Trp Lys Arg His Phe Lys Asp Leu Glu Lys Gly Cys His Ser
                965                 970                 975

Ser Ile Lys Leu Gln Arg Ser Phe Asn Lys His Gly Asn Val Phe Glu
            980                 985                 990

Cys Ser Ile Leu Glu Glu Ile Pro Tyr Glu Lys Asp Leu Ile Ile Glu
        995                 1000                1005

Arg Glu Asn Phe Trp Ile Lys Glu Leu Asn Ser Lys Ile Asn Gly Tyr
    1010                1015                1020
```

Asn Ile Ala Asp Ala Thr Phe Gly Asp Thr Cys Ser Thr His Pro Leu
1025 1030 1035 1040

Lys Glu Glu Ile Ile Lys Lys Arg Ser Glu Thr Val Lys Ala Lys Met
1045 1050 1055

Leu Lys Leu Gly Pro Asp Gly Arg Lys Ala Leu Tyr Ser Lys Pro Gly
1060 1065 1070

Ser Lys Asn Gly Arg Trp Asn Pro Glu Thr His Lys Phe Cys Lys Cys
1075 1080 1085

Gly Val Arg Ile Gln Thr Ser Ala Tyr Thr Cys Ser Lys Cys Arg Asn
1090 1095 1100

Arg Ser Gly Glu Asn Asn Ser Phe Phe Asn His Lys His Ser Gln Gly
1105 1110 1115 1120

Pro Ser Ala Asp

<210> 451

<211> 659
<212> PRT
<213> Artificial sequence

<220>
<223> scTrex2_cT11

<400> 451
Met Gly Ser Glu Ala Pro Arg Ala Glu Thr Phe Val Phe Leu Asp Leu
1 5 10 15

Glu Ala Thr Gly Leu Pro Ser Val Glu Pro Glu Ile Ala Glu Leu Ser
20 25 30

Leu Phe Ala Val His Arg Ser Ser Leu Glu Asn Pro Glu His Asp Glu
35 40 45

Ser Gly Ala Leu Val Leu Pro Arg Val Leu Asp Lys Leu Thr Leu Cys
50 55 60

Met Cys Pro Glu Arg Pro Phe Thr Ala Lys Ala Ser Glu Ile Thr Gly
65 70 75 80

Leu Ser Ser Glu Gly Leu Ala Arg Cys Arg Lys Ala Gly Phe Asp Gly
85 90 95

Ala Val Val Arg Thr Leu Gln Ala Phe Leu Ser Arg Gln Ala Gly Pro
100 105 110

```
Ile Cys Leu Val Ala His Asn Gly Phe Asp Tyr Asp Phe Pro Leu Leu
        115                 120                 125

Cys Ala Glu Leu Arg Arg Leu Gly Ala Arg Leu Pro Arg Asp Thr Val
    130                 135                 140

Cys Leu Asp Thr Leu Pro Ala Leu Arg Gly Leu Asp Arg Ala His Ser
145                 150                 155                 160

His Gly Thr Arg Ala Arg Gly Arg Gln Gly Tyr Ser Leu Gly Ser Leu
                165                 170                 175

Phe His Arg Tyr Phe Arg Ala Glu Pro Ser Ala Ala His Ser Ala Glu
            180                 185                 190

Gly Asp Val His Thr Leu Leu Leu Ile Phe Leu His Arg Ala Ala Glu
        195                 200                 205

Leu Leu Ala Trp Ala Asp Glu Gln Ala Arg Gly Trp Ala His Ile Glu
    210                 215                 220

Pro Met Tyr Leu Pro Pro Asp Asp Pro Ser Leu Glu Ala Thr Pro Pro
225                 230                 235                 240

Gln Thr Gly Leu Asp Val Pro Tyr Ser Glu Ala Pro Arg Ala Glu Thr
                245                 250                 255

Phe Val Phe Leu Asp Leu Glu Ala Thr Gly Leu Pro Ser Val Glu Pro
            260                 265                 270

Glu Ile Ala Glu Leu Ser Leu Phe Ala Val His Arg Ser Ser Leu Glu
        275                 280                 285

Asn Pro Glu His Asp Glu Ser Gly Ala Leu Val Leu Pro Arg Val Leu
    290                 295                 300

Asp Lys Leu Thr Leu Cys Met Cys Pro Glu Arg Pro Phe Thr Ala Lys
305                 310                 315                 320

Ala Ser Glu Ile Thr Gly Leu Ser Ser Glu Gly Leu Ala Arg Cys Arg
                325                 330                 335

Lys Ala Gly Phe Asp Gly Ala Val Val Arg Thr Leu Gln Ala Phe Leu
            340                 345                 350

Ser Arg Gln Ala Gly Pro Ile Cys Leu Val Ala His Asn Gly Phe Asp
        355                 360                 365

Tyr Asp Phe Pro Leu Leu Cys Ala Glu Leu Arg Arg Leu Gly Ala Arg
    370                 375                 380
```

```
Leu Pro Arg Asp Thr Val Cys Leu Asp Thr Leu Pro Ala Leu Arg Gly
385                 390                 395                 400

Leu Asp Arg Ala His Ser His Gly Thr Arg Ala Arg Gly Arg Gln Gly
                405                 410                 415

Tyr Ser Leu Gly Ser Leu Phe His Arg Tyr Phe Arg Ala Glu Pro Ser
            420                 425                 430

Ala Ala His Ser Ala Glu Gly Asp Val His Thr Leu Leu Leu Ile Phe
        435                 440                 445

Leu His Arg Ala Ala Glu Leu Leu Ala Trp Ala Asp Glu Gln Ala Arg
    450                 455                 460

Gly Trp Ala His Ile Glu Pro Met Tyr Leu Pro Pro Asp Asp Pro Ser
465                 470                 475                 480

Leu Glu Ala Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser Gly Val
                485                 490                 495

Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile Lys
            500                 505                 510

Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly
        515                 520                 525

His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala
    530                 535                 540

Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu
545                 550                 555                 560

Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser
                565                 570                 575

Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg
            580                 585                 590

Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys
        595                 600                 605

Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn Ala
    610                 615                 620

Leu Thr Gly Ala Pro Leu Asn Leu Thr Gly Asp Ala Arg Gly Ile Arg
625                 630                 635                 640

Ser Arg Ala Ser Arg Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro
                645                 650                 655
```

Ser Ala Asp

<210> 452

<211> 1452
<212> PRT
<213> Artificial sequence

<220>
<223> scTrex2_cT11Avr_FokI-L

<400> 452
Met Gly Ser Glu Ala Pro Arg Ala Glu Thr Phe Val Phe Leu Asp Leu
1   5   10   15

Glu Ala Thr Gly Leu Pro Ser Val Glu Pro Glu Ile Ala Glu Leu Ser
            20   25   30

Leu Phe Ala Val His Arg Ser Ser Leu Glu Asn Pro Glu His Asp Glu
        35   40   45

Ser Gly Ala Leu Val Leu Pro Arg Val Leu Asp Lys Leu Thr Leu Cys
    50   55   60

Met Cys Pro Glu Arg Pro Phe Thr Ala Lys Ala Ser Glu Ile Thr Gly
65   70   75   80

Leu Ser Ser Glu Gly Leu Ala Arg Cys Arg Lys Ala Gly Phe Asp Gly
                85   90   95

Ala Val Val Arg Thr Leu Gln Ala Phe Leu Ser Arg Gln Ala Gly Pro
            100   105   110

Ile Cys Leu Val Ala His Asn Gly Phe Asp Tyr Asp Phe Pro Leu Leu
        115   120   125

Cys Ala Glu Leu Arg Arg Leu Gly Ala Arg Leu Pro Arg Asp Thr Val
    130   135   140

Cys Leu Asp Thr Leu Pro Ala Leu Arg Gly Leu Asp Arg Ala His Ser
145   150   155   160

His Gly Thr Arg Ala Arg Gly Arg Gln Gly Tyr Ser Leu Gly Ser Leu
                165   170   175

Phe His Arg Tyr Phe Arg Ala Glu Pro Ser Ala Ala His Ser Ala Glu
            180   185   190

Gly Asp Val His Thr Leu Leu Leu Ile Phe Leu His Arg Ala Ala Glu
        195   200   205

Leu Leu Ala Trp Ala Asp Glu Gln Ala Arg Gly Trp Ala His Ile Glu
    210   215   220

Pro Met Tyr Leu Pro Pro Asp Asp Pro Ser Leu Glu Ala Thr Pro Pro
225 230 235 240

Gln Thr Gly Leu Asp Val Pro Tyr Ser Glu Ala Pro Arg Ala Glu Thr
245 250 255

Phe Val Phe Leu Asp Leu Glu Ala Thr Gly Leu Pro Ser Val Glu Pro
260 265 270

Glu Ile Ala Glu Leu Ser Leu Phe Ala Val His Arg Ser Ser Leu Glu
275 280 285

Asn Pro Glu His Asp Glu Ser Gly Ala Leu Val Leu Pro Arg Val Leu
290 295 300

Asp Lys Leu Thr Leu Cys Met Cys Pro Glu Arg Pro Phe Thr Ala Lys
305 310 315 320

Ala Ser Glu Ile Thr Gly Leu Ser Ser Glu Gly Leu Ala Arg Cys Arg
325 330 335

Lys Ala Gly Phe Asp Gly Ala Val Val Arg Thr Leu Gln Ala Phe Leu
340 345 350

Ser Arg Gln Ala Gly Pro Ile Cys Leu Val Ala His Asn Gly Phe Asp
355 360 365

Tyr Asp Phe Pro Leu Leu Cys Ala Glu Leu Arg Arg Leu Gly Ala Arg
370 375 380

Leu Pro Arg Asp Thr Val Cys Leu Asp Thr Leu Pro Ala Leu Arg Gly
385 390 395 400

Leu Asp Arg Ala His Ser His Gly Thr Arg Ala Arg Gly Arg Gln Gly
405 410 415

Tyr Ser Leu Gly Ser Leu Phe His Arg Tyr Phe Arg Ala Glu Pro Ser
420 425 430

Ala Ala His Ser Ala Glu Gly Asp Val His Thr Leu Leu Leu Ile Phe
435 440 445

Leu His Arg Ala Ala Glu Leu Leu Ala Trp Ala Asp Glu Gln Ala Arg
450 455 460

Gly Trp Ala His Ile Glu Pro Met Tyr Leu Pro Pro Asp Asp Pro Ser
465 470 475 480

Leu Glu Ala Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser Gly Val
485 490 495

Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Glu Lys Ile Lys
500 505 510

Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val Gly
515 520 525

His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro Ala
530 535 540

Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu
545 550 555 560

Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser
565 570 575

Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg
580 585 590

Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys
595 600 605

Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg Asn Ala
610 615 620

Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val Ala Ile
625 630 635 640

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
645 650 655

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val
660 665 670

Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
675 680 685

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
690 695 700

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
705 710 715 720

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
725 730 735

Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu
740 745 750

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
755 760 765

```
Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
    770                 775                 780

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
785                 790                 795                 800

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
                805                 810                 815

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
            820                 825                 830

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile
        835                 840                 845

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu
    850                 855                 860

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
865                 870                 875                 880

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
                885                 890                 895

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
            900                 905                 910

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
        915                 920                 925

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val
    930                 935                 940

Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
945                 950                 955                 960

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                965                 970                 975

Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr
            980                 985                 990

Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        995                 1000                1005

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
    1010                1015                1020

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
1025                1030                1035                1040
```

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
1045 1050 1055

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
1060 1065 1070

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
1075 1080 1085

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
1090 1095 1100

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
1105 1110 1115 1120

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
1125 1130 1135

Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser
1140 1145 1150

Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
1155 1160 1165

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
1170 1175 1180

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
1185 1190 1195 1200

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val
1205 1210 1215

Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val
1220 1225 1230

Ala Gln Leu Ser Arg Pro Asp Pro Gly Ser Ser Gly Pro Asn Arg Gly
1235 1240 1245

Val Thr Lys Gln Leu Val Lys Ser Glu Leu Glu Glu Lys Lys Ser Glu
1250 1255 1260

Leu Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile
1265 1270 1275 1280

Glu Ile Ala Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val
1285 1290 1295

Met Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly
1300 1305 1310

Gly Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile
1315 1320 1325

Asp Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn
1330 1335 1340

Leu Pro Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn
1345 1350 1355 1360

Gln Thr Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr
1365 1370 1375

Pro Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe
1380 1385 1390

Lys Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn
1395 1400 1405

Cys Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu
1410 1415 1420

Met Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe
1425 1430 1435 1440

Asn Asn Gly Glu Ile Asn Phe Gly Ser Ser Ala Asp
1445 1450

<210> 453

<211> 1570
<212> PRT
<213> Artificial sequence

<220>
<223> cT11Avr_ColE7_A497_cT11RagT2-R

<400> 453
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

```
Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
                85                  90                  95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
            100                 105                 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Glu Gln Val Val
    130                 135                 140

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145                 150                 155                 160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
                165                 170                 175

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
            180                 185                 190

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        195                 200                 205

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
    210                 215                 220

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225                 230                 235                 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
                245                 250                 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            260                 265                 270

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
        275                 280                 285

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
    290                 295                 300

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile
305                 310                 315                 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu
                325                 330                 335

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
            340                 345                 350
```

```
Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
        355                 360                 365

Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
    370                 375                 380

Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
385                 390                 395                 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
                405                 410                 415

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
            420                 425                 430

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
        435                 440                 445

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
    450                 455                 460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465                 470                 475                 480

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
                485                 490                 495

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
            500                 505                 510

Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln
        515                 520                 525

Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His
    530                 535                 540

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
545                 550                 555                 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
                565                 570                 575

Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp
            580                 585                 590

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
        595                 600                 605

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
    610                 615                 620
```

```
His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
625                 630                 635                 640

Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
                645                 650                 655

Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
            660                 665                 670

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val
        675                 680                 685

Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
    690                 695                 700

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
705                 710                 715                 720

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser
                725                 730                 735

Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Gly Ser Ser Gly Ser Lys
            740                 745                 750

Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys Gly Lys Pro Val Asn Asn
        755                 760                 765

Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu Gly Ser Pro Val Pro Asp
    770                 775                 780

Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu Phe Lys Ser Phe Asp Asp
785                 790                 795                 800

Phe Arg Ala Lys Phe Trp Glu Glu Val Ser Lys Asp Pro Glu Leu Ser
                805                 810                 815

Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg Met Lys Val Gly Lys Ala
            820                 825                 830

Pro Lys Thr Arg Thr Gln Asp Val Ser Gly Lys Arg Thr Ser Phe Glu
        835                 840                 845

Leu His His Glu Lys Pro Ile Ser Gln Asn Gly Gly Val Tyr Asp Met
    850                 855                 860

Asp Asn Ile Ser Val Val Thr Pro Lys Arg His Ile Asp Ile His Arg
865                 870                 875                 880

Gly Lys Gly Ser Thr Gly Gly Ser Ser Gly Val Asp Leu Arg Thr Leu
                885                 890                 895
```

Gly Tyr Ser Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser
900 905 910

Thr Val Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His
915 920 925

Ala His Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val
930 935 940

Ala Val Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His
945 950 955 960

Glu Ala Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu
965 970 975

Glu Ala Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln
980 985 990

Leu Asp Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr
995 1000 1005

Ala Val Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro
1010 1015 1020

Leu Asn Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly
1025 1030 1035 1040

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
1045 1050 1055

Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
1060 1065 1070

Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
1075 1080 1085

Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala
1090 1095 1100

Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
1105 1110 1115 1120

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala
1125 1130 1135

Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
1140 1145 1150

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
1155 1160 1165

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val
1170 1175 1180

Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln
1185 1190 1195 1200

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu
1205 1210 1215

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
1220 1225 1230

Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala
1235 1240 1245

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
1250 1255 1260

Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys
1265 1270 1275 1280

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
1285 1290 1295

His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly
1300 1305 1310

Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys
1315 1320 1325

Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
1330 1335 1340

Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
1345 1350 1355 1360

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
1365 1370 1375

Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu
1380 1385 1390

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
1395 1400 1405

Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
1410 1415 1420

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val
1425 1430 1435 1440

Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val
1445 1450 1455

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln
1460 1465 1470

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu
1475 1480 1485

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
1490 1495 1500

Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala
1505 1510 1515 1520

Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly
1525 1530 1535

Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg
1540 1545 1550

Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ser
1555 1560 1565

Ala Asp
1570

<210> 454

<211> 1570
<212> PRT
<213> Artificial sequence

<220>
<223> cT11RagT2-R_ColE7_A497_cT11Avr

<400> 454
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
100 105 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
115 120 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Gln Gln Val Val
130 135 140

Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145 150 155 160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
165 170 175

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
180 185 190

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
195 200 205

Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu
210 215 220

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225 230 235 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln
245 250 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
260 265 270

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly
275 280 285

Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln
290 295 300

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
305 310 315 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
325 330 335

Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser
340 345 350

Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
355 360 365

```
Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
    370                 375                 380

Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu
385                 390                 395                 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val
                405                 410                 415

Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
            420                 425                 430

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln
        435                 440                 445

Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr
    450                 455                 460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465                 470                 475                 480

Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu
                485                 490                 495

Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
            500                 505                 510

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
        515                 520                 525

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
    530                 535                 540

Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly
545                 550                 555                 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
                565                 570                 575

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
            580                 585                 590

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
        595                 600                 605

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
    610                 615                 620

Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro
625                 630                 635                 640
```

```
Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile
                645                 650                 655

Ala Ser Asn Gly Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln
            660                 665                 670

Leu Ser Arg Pro Asp Pro Gly Ser Ser Gly Ser Lys Arg Asn Lys Pro
        675                 680                 685

Gly Lys Ala Thr Gly Lys Gly Lys Pro Val Asn Asn Lys Trp Leu Asn
    690                 695                 700

Asn Ala Gly Lys Asp Leu Gly Ser Pro Val Pro Asp Arg Ile Ala Asn
705                 710                 715                 720

Lys Leu Arg Asp Lys Glu Phe Lys Ser Phe Asp Asp Phe Arg Ala Lys
                725                 730                 735

Phe Trp Glu Glu Val Ser Lys Asp Pro Glu Leu Ser Lys Gln Phe Ser
            740                 745                 750

Arg Asn Asn Asn Asp Arg Met Lys Val Gly Lys Ala Pro Lys Thr Arg
        755                 760                 765

Thr Gln Asp Val Ser Gly Lys Arg Thr Ser Phe Glu Leu His His Glu
    770                 775                 780

Lys Pro Ile Ser Gln Asn Gly Gly Val Tyr Asp Met Asp Asn Ile Ser
785                 790                 795                 800

Val Val Thr Pro Lys Arg His Ile Asp Ile His Arg Gly Lys Gly Ser
                805                 810                 815

Thr Gly Gly Ser Ser Gly Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln
            820                 825                 830

Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln
        835                 840                 845

His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile Val
    850                 855                 860

Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr
865                 870                 875                 880

Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val
                885                 890                 895

Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu
            900                 905                 910
```

Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly
915 920 925

Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala
930 935 940

Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr
945 950 955 960

Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala
965 970 975

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
980 985 990

Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys
995 1000 1005

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
1010 1015 1020

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly
1025 1030 1035 1040

Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys
1045 1050 1055

Gln Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn
1060 1065 1070

Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val
1075 1080 1085

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
1090 1095 1100

Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu
1105 1110 1115 1120

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
1125 1130 1135

Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala
1140 1145 1150

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
1155 1160 1165

Val Ala Ile Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val
1170 1175 1180

```
Gln Ala Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu
1185                1190                1195                1200

Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu
                1205                1210                1215

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
            1220                1225                1230

Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala
        1235                1240                1245

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
    1250                1255                1260

Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys
1265                1270                1275                1280

Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
                1285                1290                1295

His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Ile Gly
            1300                1305                1310

Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val Leu Cys
        1315                1320                1325

Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn
    1330                1335                1340

Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu Leu Pro Val
1345                1350                1355                1360

Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala
                1365                1370                1375

Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu
            1380                1385                1390

Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala
        1395                1400                1405

Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
    1410                1415                1420

Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val
1425                1430                1435                1440

Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val
                1445                1450                1455
```

Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln
1460 1465 1470

Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu
1475 1480 1485

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr
1490 1495 1500

Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala
1505 1510 1515 1520

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly
1525 1530 1535

Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg
1540 1545 1550

Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ser
1555 1560 1565

Ala Asp
1570

<210> 455

<211> 977
<212> PRT
<213> Artificial sequence

<220>
<223> cT11TB2A2_scTB2aD01_148C

<400> 455
Met Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys
1 5 10 15

Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu
20 25 30

Val Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His
35 40 45

Pro Ala Ala Leu Gly Thr Val Ala Val Lys Tyr Gln Asp Met Ile Ala
50 55 60

Ala Leu Pro Glu Ala Thr His Glu Ala Ile Val Gly Val Gly Lys Gln
65 70 75 80

Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Val Ala Gly Glu
85 90 95

```
Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Leu Lys Ile
            100                 105                 110

Ala Lys Arg Gly Gly Val Thr Ala Val Glu Ala Val His Ala Trp Arg
        115                 120                 125

Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Gln Gln Val Val
    130                 135                 140

Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
145                 150                 155                 160

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
                165                 170                 175

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
            180                 185                 190

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
        195                 200                 205

Gln Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu
    210                 215                 220

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
225                 230                 235                 240

Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln
                245                 250                 255

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
            260                 265                 270

Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly
        275                 280                 285

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
    290                 295                 300

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
305                 310                 315                 320

Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu
                325                 330                 335

Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser
            340                 345                 350

His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro
        355                 360                 365
```

```
Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln Val Val Ala Ile
    370                 375                 380

Ala Ser Asn Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Ala Leu
385                 390                 395                 400

Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Gln Gln Val Val
                405                 410                 415

Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln
            420                 425                 430

Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Glu Gln
        435                 440                 445

Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr
    450                 455                 460

Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro
465                 470                 475                 480

Gln Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu
                485                 490                 495

Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu
            500                 505                 510

Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln
        515                 520                 525

Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His
    530                 535                 540

Gly Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly
545                 550                 555                 560

Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln
                565                 570                 575

Ala His Gly Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Gly
            580                 585                 590

Gly Gly Arg Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro
        595                 600                 605

Asp Pro Gly Ser Asp Ile Thr Lys Ser Lys Ile Ser Glu Lys Met Lys
    610                 615                 620

Gly Gln Gly Pro Ser Gly Asn Thr Lys Tyr Asn Glu Glu Phe Leu Leu
625                 630                 635                 640
```

Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile Val Ala Gln Ile
645 650 655

Lys Pro His Gln Ser Cys Lys Phe Lys His Gln Leu Arg Leu Thr Phe
660 665 670

Asn Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val
675 680 685

Asp Glu Ile Gly Val Gly His Val Tyr Asp Ser Gly Ser Val Ser Tyr
690 695 700

Tyr Gln Leu Ser Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu
705 710 715 720

Gln Pro Phe Leu Glu Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys
725 730 735

Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu
740 745 750

Glu Val Cys Thr Trp Val Asp Gln Val Ala Ala Leu Asn Asp Ser Lys
755 760 765

Thr Arg Lys Thr Thr Ser Glu Thr Cys Arg Ala Val Leu Asp Ser Leu
770 775 780

Ser Glu Lys Lys Lys Ser Ser Pro Ala Ala Gly Asp Ser Ser Val Ser
785 790 795 800

Asn Ser Glu His Ile Ala Pro Leu Ser Leu Pro Ser Ser Pro Pro Ser
805 810 815

Val Gly Ser Asn Lys Lys Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp
820 825 830

Ser Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys
835 840 845

Phe Lys His Tyr Leu Ser Leu Thr Phe Cys Val Thr Gln Lys Thr Gln
850 855 860

Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Arg Ile Gly Val Gly Tyr
865 870 875 880

Val Gln Asp Arg Gly Ser Val Ser Asp Tyr Arg Leu Ser Glu Ile Lys
885 890 895

Pro Leu Arg Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys
900 905 910

```
Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser
        915                 920                 925

Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp
    930                 935                 940

Gln Val Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu
945                 950                 955                 960

Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser
                965                 970                 975

Pro
```