

APPLICATION FOR PATENT

FOR

NUCLEOTIDE SEQUENCES ENCODING INSECTICIDAL PROTEINS

Natalia N. Bogdanova

David R. Corbin

Thomas M. Malvar

Frederick J. Perlak

James K. Roberts

Charles P. Romano

BACKGROUND OF THE INVENTION

The present invention provides novel coding sequences for use in plants. The coding sequences encode a chimeric insecticidal protein toxic to a wide range of lepidopteran species crop pests.

Commercial formulations of naturally occurring *B. thuringiensis* isolates have long been used for the biological control of agricultural insect pests. Bt spores and crystals obtained from fermentation of *Bacillus thuringiensis* species are concentrated and formulated for foliar application according to conventional agricultural practices.

Members of the family of Cry1 crystal proteins are known to exhibit bioactivity against lepidopteran insect larvae and are useful as agents for controlling lepidopteran insect pests. The precursor form of Cry1 δ -endotoxins consist of two approximately equal sized segments. The carboxy-terminal portion of the precursor protein, or pro-toxin segment, stabilizes crystal formation and exhibits no insecticidal activity. The amino-terminal half of the precursor protein comprises the toxin segment of the Cry1 protein and, based on alignment of conserved or substantially conserved sequences within Cry1 family members, can be further sub-divided into three structural domains. These three sub-domains are based on a three dimensional crystallographic structural model of a Cry1A δ -endotoxin in which the three sub-domains were referred to as Domain I, Domain II, and Domain III, respectively as measured from the amino terminus of the protein toxin segment. Domain I comprises about the first third of the active toxin segment and has been shown to be essential for channel formation (Thompson et al., 1995). Domains II and III respectively comprise about the central and carboxy-terminal segments of the active toxin portion. Domains II and III have both been implicated in receptor binding and insect species specificity, depending on the insect and δ -endotoxin being examined (Thompson et al., 1995).

The likelihood of arbitrarily creating a chimeric protein with enhanced properties from the reassortment of the domain structures of the numerous native insecticidal crystal proteins known in the art is remote. This is a result of the complex nature of protein structure, folding, oligomerization, and activation including correct proteolytic processing of the chimeric precursor, if expressed in such form, to release an insecticidal toxin segment. Only by careful selection of specific target regions within each parental protein for inclusion into a chimeric structure can functional insecticidal toxins be constructed that exhibit improved insecticidal activity in comparison to the parental proteins from which the chimeras are derived. Experience has shown that reassembly of the toxin domains, i.e., assembly of a

chimeric toxin consisting of domain I, II, and III of any two or more toxins that are different from each other, results in the construction of a protein that exhibits faulty crystal formation and/or the complete lack of any detectable insecticidal activity directed to a preferred target insect pest species. In some instances, a chimeric toxin will exhibit good crystal formation properties, yet exhibit no detectable insecticidal activity. Only by trial and error are effective insecticidal chimeras formulated, and even then, the skilled artisan is not certain to end up with a chimera that exhibits insecticidal activity that is equivalent to or improved in comparison to any single parental toxin protein from which the constituents of the chimera may have been derived.

The literature reports examples of the construction or assembly of chimeric proteins from two or more Bt insecticidal crystal protein precursors, yet not all exhibited insecticidal or crystal forming properties that were equivalent to or improved in comparison to the precursor proteins from which the chimeras were derived. (Bosch et al. (WO95/06730); Thompson et al. (WO95/30753); Thompson et al. (WO95/30752); Malvar et al. (WO98/22595); Gilroy et al. (US Patent No. 5,128,130); Gilroy (US Patent No. 5,055,294); Lee et al. (1992) Gene 267:3115-3121; Honee et al. (1991) Mol. Microbiol. 5 :2799-2806; Schnepf et al. (1990) J. Biol. Chem. 265:20923-20930; Perlak et al. (1990) Bio/Technol. 8:939-9943; Perlak et al (1993) Plant Mol. Biol. 22:313-321).

Expression of *B. thuringiensis* δ -endotoxins in transgenic corn plants has proven to be an effective means of controlling agriculturally important insect pests (Perlak et al. 1990; 1993). Transgenic crops expressing *B. thuringiensis* δ -endotoxins enable growers to significantly reduce the time and costs associated with applications of topically applied chemical insecticides. Use of transgenes encoding *B. thuringiensis* δ -endotoxins is particularly advantageous. Crop plants expressing *B. thuringiensis* δ -endotoxins in areas under heavy insect pressure exhibit improved yields that are better than otherwise similar non-transgenic commercial plant varieties. However, it is anticipated that insects may evolve resistance to *B. thuringiensis* δ -endotoxins expressed in transgenic plants. Such resistance, should it become widespread, would clearly limit the commercial value of germplasm containing genes encoding such *B. thuringiensis* δ -endotoxins. One possible way of increasing the effectiveness of the transgenic insecticides against target pests and contemporaneously reducing the development of insecticide-resistant pests would be to ensure that transgenic crops express high levels of *B. thuringiensis* δ -endotoxins (McGaughey and Whalon 1993; Roush 1994). In addition, having a repository of

insecticidal genes that are effective against groups of insect pests and which manifest their effects through different modes of action can safeguard against any development of resistance. Expression in a plant of two or more insecticidal compositions toxic to the same insect species, each insecticide being expressed at levels high enough to effectively delay the onset of resistance, would be another way to achieve control of the development of resistance. Examples of such insecticides useful in such combinations include but are not limited to Bt toxins, *Xenorhabdus* sp. or *Photorhabdus* sp. insecticidal proteins, deallergized and de-glycosylated patatin proteins and/or permuteins, plant lectins, and the like. Achieving co-expression of multiple insecticidally active proteins in the same plant, and/or high expression levels of those insecticidal proteins without causing undesirable plant morphological effects has been elusive.

Only a handful of the more than two-hundred and fifty individual insecticidal proteins that have been identified from *Bacillus thuringiensis* species have been tested for expression in plants. Several Cry1's, Cry3's, Cry2Aa and Cry2Ab, binary toxins Cry33/34 and Cry23/37, and a Cry9 have been successfully expressed in plants. Cry1 proteins represent the largest class of proteins that have been expressed in plants, but none have been expressed at high levels. It was necessary to target the Cry2Ab to the chloroplast to avoid undesirable phytotoxic effects. The majority of acres planted in recombinant plants express Cry1A proteins. The likelihood of the onset of resistance to Cry1A proteins by targeted insect pest species is substantially higher than it would be if a resistance management allele was also expressed along with the *cry1* allele, or if the *cry1* allele was expressed at high levels. Therefore it is desirable that alternative toxin genes be developed for expression in plants as supplements and replacements for those being used presently in the first and second generations of transgenic insect resistant plants.

SUMMARY OF THE INVENTION

The invention provides isolated nucleotide sequences for expression in plants encoding an insecticidal protein exhibiting lepidopteran insect inhibitory properties. SEQ ID NO:1 is an example of such nucleotide sequences consisting of a *cry1A.105* gene and encodes an insect inhibitory Cry1A.105 protein. SEQ ID NO:1 is similar to SEQ ID NO:3, both encoding a Cry1A.105 protein. SEQ ID NO:1 is preferred for use in a dicotyledonous cells, while SEQ ID NO:3 is preferred for use in monocotyledonous cells. SEQ ID NO:4 is encoded from SEQ ID NO:3 and is identical in amino acid sequence to SEQ ID NO:2. The isolated nucleotide sequence is intended to include sequences that exhibit at least from about 88% to about 90% or greater nucleotide sequence identity to the sequence as set forth at SEQ ID NO:1, or that hybridize to SEQ ID NO:1 under stringent hybridization conditions. The isolated nucleotide sequence is also intended to include sequences that exhibit at least about 90% nucleotide sequence identity to the sequence as set forth at SEQ ID NO:3, or that hybridize to SEQ ID NO:3 under stringent hybridization conditions.

The invention also provides an isolated and purified insecticidal protein exhibiting inhibitory activity directed to lepidopteran insect species. The insecticidal protein is designated herein at least as the toxin portion of Cry1A.105 and exhibits an amino acid sequence as set forth in SEQ ID NO:2. The full length precursor protein consisting of about 1177 amino acids as set forth in SEQ ID NO:2 is also referred to as an insecticidal Cry1A.105 protein, however any fragment of the precursor protein that exhibits insecticidal bioactivity is intended to be referred to as an insecticidal Cry1A.105 protein, and includes at least a Cry1A.105 insecticidal protein corresponding to an amino acid sequence segment from about amino acid 1 through about amino acid 612 as set forth in SEQ ID NO:2, and may also include a segment from about amino acid 2 through about amino acid 610. Any composition consisting of an insecticidally effective amount of the insecticidal protein is intended to be within the scope of the invention.

The invention also provides an expression cassette for use in expressing an insecticidal protein as set forth in SEQ ID NO:2 in a host cell. The expression cassette preferably contains a promoter functional in the intended host cell which is linked to and regulates the expression of a nucleotide sequence encoding an insecticidal segment of a Cry1A.105 protein. Exemplary expression cassettes are provided herein as set forth at SEQ ID NO:5 and SEQ ID NO:7, intended for use in a dicot plant cell or a monocot plant cell, respectively. The promoter and the coding sequence are operably linked and function

together in the host cell. The expression cassette can be intended for use in any host cell, but is preferably for use in a bacterial cell, a fungal cell, a mammalian cell, or a plant cell. Bacterial cells are preferably selected from the group consisting of a *Bacillus* species cell, a *Enterobacteriaceae* species cell, a *Pseudomonas* species cell, a *Clostridium* species cell, and a *Rhizobium* species cell, and a *Agrobacterium* species cell. If the host cell is a plant cell, it is preferable that it is a cell chosen from a crop species of plant cell, preferably either a dicotyledonous plant or a monocotyledonous plant cell. Examples of dicotyledonous plant cells are alfalfa, apple, apricot, asparagus, bean, berry, blackberry, blueberry, canola, carrot, cauliflower, celery, cherry, chickpea, citrus tree, cotton, cowpea, cranberry, cucumber, cucurbit, egg plant, fruit tree, grape, lemon, lettuce, linseed, melon, mustard, nut bearing tree, okra, orange, pea, peach, peanut, pear, plum, potato, soybeans, squash, strawberry, sugar beet, sunflower, sweet potato, tobacco, tomato, turnip, and vegetable. Monocotyledonous plant cells examples are corn, wheat, oat, rice, sorghum, milo, buckwheat, rye, grass (fescue, timothy, brome, orchard, St. Augustine, Bermuda, bentgrass), and barley. Expression cassettes intended for use in a plant cell typically contain in operable linkage sequences that regulate the levels and efficiencies of expression of an intended substance, such as a Cry1A.105 insecticidal protein. Such sequences may be an expression enhancer sequence, an untranslated leader sequence, an intron sequence, a chloroplast targeting peptide encoding sequence, and a transcription termination and polyadenylation sequence.

The expression cassette is preferably incorporated into a vector for use in stabilizing the maintenance of the Cry1A.105 coding sequence within the host cell. A vector can be any number of structures known in the art, but is typically a plasmid or replicon into which the expression cassette is constructed or inserted prior to incorporation into the host cell. A vector is intended to include but not be limited to a plasmid, a cosmid, a bacmid, a phagemid, a YAC, a BAC, a suicide vector, an insertion sequence, a transposon, or even a linear nucleotide sequence to which the expression cassette is linked or in which the expression cassette is embedded.

Transgenic plants resistant to lepidopteran insect infestation are an embodiment of the present invention. Such plants contain a nucleotide sequence that encodes a Cry1A.105 insecticidal protein as set forth in SEQ ID NO:2 at least from about amino acid 2 to about amino acid 612. The transgenic plant is effective in controlling lepidopteran insect infestations brought about by insects such as leaf rollers, cutworms, armyworms, borers, bagworms, and any forage feeder. Preferred pests are fall armyworms, European corn borers, corn earworms (cotton bollworms), southwestern corn borers, and black cutworms.

The present invention is intended to include the progeny and seed or fruits or product yielded from the transgenic plant of the present invention, so long as the nucleotide sequence of the present invention encoding a Cry1A.105 insecticidal segment is maintained within the heritable and/or plastid genome of the cells of the plant, its progeny, seed, and the like.

5 The present invention also provides one or more methods for controlling lepidopteran insect infestation of a plant by providing in the diet of an insect pest a composition that contains an insecticidally effective amount of an insecticidal Cry1A.105 protein. One such composition would be plant cells that have been or are descended from a plant cell transformed with a nucleic acid sequence that encodes an insecticidal segment of a
10 Cry1A.105 amino acid sequence as set forth in SEQ ID NO:2. A transgenic plant generated from a plant cell transformed to contain an expression cassette, exemplified as set forth at SEQ ID NO:5 and SEQ ID NO:7, which contains a sequence encoding a Cry1A.105 insecticidal amino acid sequence, would be one means for providing an insecticidal composition in the diet of the insect. Another means would be to produce an insecticidally
15 effective amount of a Cry1A.105 protein in a bacterial or fungal cell and provide the bacterial cell or fungal cell or a purified amount of the Cry1A.105 protein in the diet of one or more target insect pests susceptible to the Cry1A.105 protein.

A method of identifying a nucleotide sequence encoding a Cry1A.105 amino acid sequence in a biological sample is provided. The method consists of contacting a sample
20 being tested for the presence of the Cry1A.105 coding sequence with a polynucleotide probe that binds with specificity to the Cry1A.105 coding sequence. In particular, the probe sequence binds, or hybridizes to, a Cry1A.105 coding sequence under stringent hybridization conditions. Detecting binding in a reaction mix is diagnostic for the presence of the Cry1A.105 coding sequence.

25 A method of identifying an insecticidal fragment of a Cry1A.105 protein in a sample is also provided. The method consists of contacting a sample being tested for the presence of a Cry1A.105 insecticidal fragment with an antibody that binds specifically to the insecticidal fragment. Detecting the binding in a reaction mix is diagnostic for the presence of the Cry1A.105 protein in the sample.

30 Chimeric or hybrid insecticidal proteins are also provided. Such hybrids are composed of two or more different insecticidal proteins, each of which exhibits insecticidal activity directed to at least one member of the same insect species. The hybrid insecticidal protein is made up of parts of each of the different insecticidal proteins. Segments of insecticidal proteins used in constructing the hybrid consist of from at least about 50 to at

least about 200 contiguous amino acids selected from the contiguous amino acids making up any one of the different insecticidal proteins. A Cry1A.105 insecticidal protein as set forth in SEQ ID NO:2 from about amino acid position 2 through about amino acid position 612 is intended to be included within the group of the different insecticidal proteins from which a
5 segment may be selected for constructing a hybrid insecticidal protein.

Various advantages and features of the present invention being apparent, the nature of the invention may be more clearly understood by reference to the following detailed description, the examples, and to the appended claims.

Brief Description of the Sequences

SEQ ID NO:1 is a synthetic sequence for expression of a Cry1A.105 insecticidal protein, preferably in a dicot cell.

5 SEQ ID NO:2 is a Cry1A.105 protein encoded from the nucleotide sequence as set forth at SEQ ID NO:1.

SEQ ID NO:3 is a synthetic sequence for expression of a Cry1A.105 insecticidal protein, preferably in a monocot cell.

SEQ ID NO:4 is a Cry1A.105 protein encoded from the nucleotide sequence as set forth at SEQ ID NO:3.

10 SEQ ID NO:5 represents a nucleotide sequence consisting of an expression cassette that functions in a plant cell, and preferably in a dicot plant cell, for expressing a Cry1A.105 insecticidal protein.

SEQ ID NO:6 represents a Cry1A.105 insecticidal protein encoded by a segment within the expression cassette as set forth in SEQ ID NO:5.

15 SEQ ID NO:7 represents a nucleotide sequence consisting of an expression cassette that functions in a plant cell, and preferably in a monocot plant cell, for expressing a Cry1A.105 insecticidal protein.

SEQ ID NO:8 represents a Cry1A.105 insecticidal protein encoded by a segment within the expression cassette as set forth in SEQ ID NO:7.

Detailed Description of the Invention

In accordance with the present invention, the inventors have constructed nucleotide sequences that encode a novel insecticidal protein identified herein as a Cry1A.105 protein. It has been identified that the Cry1A.105 amino acid sequence, set forth in SEQ ID NO:2, exhibits properties that provide advantages over naturally occurring Bt insecticidal proteins that are toxic to lepidopteran insect species. In particular, the Cry1A.105 protein can be expressed at high levels in both monocot and dicot plants without most transgenic events exhibiting phytotoxic effects as a result of the increased levels of expression compared to effects observed when naturally occurring Cry1 proteins are expressed in plants. In addition, the Cry1A.105 protein form stable crystals when expressed in *Bacillus thuringiensis*, likely because of the stabilizing effect of the Cry1Ac protoxin segment linked to the toxin moiety of the chimeric Cry1A.105 protein. In addition, the Cry1A.105 insecticidal protein exhibits a range of insecticidal bioactivity directed to lepidopteran species that is not observed with other naturally occurring Cry1 proteins that have been identified to date. Therefore, expression of the Cry1A.105 protein in transgenic plants results in increased numbers of morphologically normal transgenic events expressing higher levels of an analogue of a Cry1 toxin that exhibits a broad range of control of lepidopteran insect pest species for any event that is selected for commercial development. Such events should result in the advantage of delaying the onset of resistance to the Cry1A toxin analogue, and when combined with a second toxin that is toxic to one or more of the insect pest species to which the Cry1A analogue is also toxic and that exerts its mode of action in a way that is different from that of the Cry1A analogue, any likelihood of the development of resistance to either toxin is anticipated to be extremely remote.

The inventors have constructed at least two different nucleotide sequences for use in plants, each nucleotide sequence encoding the same Cry1A.105 insecticidal protein. The first (or amino terminal) about two thirds of the insecticidal portion of the Cry1A.105 protein consists of amino acid sequences derived from a Cry1Ab amino acid sequence. This sequence is linked to the carboxy-terminus of the toxin portion and a part of the protoxin domain of an amino acid sequence derived from an insecticidal Cry1 protein obtained from an Ecogen *Bt aizawai* strain EG6346 (Chambers et al., 1991, J. Bacteriol. 173:3966-3976). The Cry1A.105 toxin segment is linked then to a segment that is substantially a Cry1Ac protoxin peptide sequence. The inventors demonstrated that this construction provides a unique amino acid sequence that exhibits surprisingly improved insecticidal properties when compared to the properties exhibited by the protein from which the chimera is derived.

Furthermore, the Cry1A.105 precursor protein exhibits excellent crystal forming properties and is efficiently solubilized and processed to the active toxin form in the gut of specific targeted lepidopteran insect pests.

The nucleotide sequences embodied herein have been constructed using methods set forth in US Patents No. 5,500,365, and 5,689,052, in particular by avoiding certain inimical sequences in the coding sequence that have been observed to be problematic for expression of heterologous gene sequences in plant cells. The segment encoding the toxin portion of the Cry1A.105 protein consists of nucleotides as set forth in SEQ ID NO:1 and SEQ ID NO:3 from about position 1 through about position 1830, more or less. The sequence as set forth at SEQ ID NO:1 was constructed for use in dicotyledonous plant species, and in particular, in cotton plants. The sequence as set forth at SEQ ID NO:3 was constructed for expression in monocotyledonous plants, and in particular, in maize or corn plant species.

Nucleotide sequences of the present invention exhibit an overall identity of about 94.3% to each other and are identical from about nucleotide position 1330 through about nucleotide position 3534. The segment of each of these nucleotide sequences encoding the toxin portion of the Cry1A.105 protein exhibits, from about nucleotide position 1 through about nucleotide position 1830, about 88.9% identity to each other. The segment of these nucleotide sequences encoding the first two domain structures of the Cry1A.105 protein is substantially more diverse and exhibits only about 84.7% identity to each other.

The inventors have constructed transgenic plant events using these sequences.

SEQ ID NO:1 was introduced into a plasmid vector containing an expression cassette consisting of a enhanced Figwort Mosaic Virus promoter (eFMV) sequence operably linked to a *Petunia hybrida* Hsp70 untranslated leader sequence (Ph.Hsp70, a.k.a., DnaK), an *Arabidopsis thaliana* ribulose bis phosphate carboxylase small subunit chloroplast targeting peptide coding sequence, and a *Pisum sativum* E9 ribulose bis phosphate carboxylase small subunit gene transcription termination and polyadenylation sequence. The Cry1A.105 coding sequence as set forth at SEQ ID NO:1 was inserted into this expression cassette in frame with and immediately adjacent to the 3' end coding sequence of the targeting peptide coding sequence, and upstream of the E9 termination sequence. The nucleotide sequence of the resulting expression cassette is set forth at SEQ ID NO:5. A segment of the vector containing the Cry1A.105 expression cassette linked to a second expression cassette containing a plant expressible GUS marker was excised and used to generate transgenic cotton events using biolistic methods. Transgenic events were tested in bioassay for insecticidal activity against several different lepidopteran pest species and were determined to

exhibit significantly better insect controlling properties than previously existing insect resistant cotton plants containing only Cry1Ac or a combination of Cry1Ac and Cry2Ab proteins. In addition, some of the Cry1A.105 transgenic cotton events exhibited levels of Cry1A.105 protein accumulation exceeding 10 to 20 parts per million throughout the growing season, even in cotton bolls, and without exhibiting any phytotoxic effects on the plant or reproductive tissues. This is in contrast to other Cry1 proteins that have been tested previously, which generally were only capable of levels of accumulation to less than about 10 parts per million, whether or not targeted to the chloroplast. Phytotoxic effects were observed when other Cry1 type proteins were tested in cotton, especially when levels of Cry1 accumulation approached or exceeded about 10 ppm.

SEQ ID NO:3 was introduced into a plasmid vector containing an expression cassette consisting of a enhanced Cauliflower Mosaic Virus promoter (eCaMV) sequence operably linked to a *Triticum aestivum* major chlorophyll a/b binding protein gene untranslated leader sequence and an *Oryza sativa* actin intron sequence, and a *Triticum aestivum* hsp17 gene transcription termination and polyadenylation sequence. The Cry1A.105 coding sequence as set forth at SEQ ID NO:3 was inserted into this expression cassette immediately adjacent to and 3' of the intron sequence and upstream of the termination sequence. The nucleotide sequence of the resulting expression cassette is set forth at SEQ ID NO:7. The vector also contains a glyphosate herbicide selectable marker that was used to select events transformed with the Cry1A.105 expression cassette. Maize events selected after transformation with the Cry1A.105 expression cassette were tested in bioassays against several lepidopteran pest species and determined to exhibit a wide range of insecticidal activity that was not prevalent with events transformed with other Bt insecticidal proteins such as Cry1Ab. The fall armyworm and black cutworm activities exhibited by events expressing insecticidal levels of Cry1A.105 coupled with the Cry1A.105 insecticidal activity directed to corn earworm and corn borer equivalent to or greater than that of events expressing Cry1Ab, provides a broader spectrum of insecticidal activity for Cry1A.105 events.

The nucleotide sequences of the present invention are exemplary. Other nucleotide sequences are capable of expressing a Cry1A.105 insecticidal protein fragment in a plant cell, and still other nucleotide sequences are capable of being designed that express well in other types of host cells. Without limiting the scope of the disclosure, it is intended that a nucleotide sequence for use in expression of a Cry1A.105 insecticidal fragment exhibit at least about 85%, or at least about 90%, or at least about 95%, or at least about 99% or greater nucleotide sequence identity to the nucleotide sequences exemplified herein. Other

nucleotide sequences intended for expression of a Cry1A.105 insecticidal fragment in a host cell other than a plant cell can be of any percentage identity or similarity to the exemplified nucleotide sequences. Nucleotide sequences can vary because of the redundancy of the genetic code, and so it is possible to synthesize any number of nucleic acid sequences that encode any part of the amino acid sequence set forth in SEQ ID NO:2, and all of these sequences are intended to be within the scope of the present invention. Any isolated and purified nucleic acid sequence encoding at least an insecticidal fragment of a Cry1.105 protein is intended to be within the scope of the disclosure, as well as any composition in which the nucleic acid can be detected by antibody, by nucleic acid probe, or by one or more pairs of primers designed to produce an amplicon consisting of such sequence.

The nucleic acid sequence exemplified herein and expressed in maize consists only of a Cry1A.105 precursor protein coding sequence, while the sequence expressed in cotton consists of a chloroplast targeted Cry1A.105 precursor protein coding sequence. The expression of Cry1 proteins in plants has proven to be problematic. It is not known whether or if any particular Cry1 protein will be expressed well in any particular plant, and so trial and error experimentation is required. Some Cry1 proteins expressed in corn will result in phytotoxic effects, and so targeting the protein to the chloroplast sometimes alleviates such effects. Similar circumstances are observed with cotton plant expression of Cry1 proteins. The examples herein are not intended to teach that Cry1A.105 expression is only possible in maize if localized to the cytoplasmic space, and similarly, are not intended to teach that Cry1A.105 expression is only possible in cotton if localized to the plastid. The examples are intended to teach that either method of protein localization functions with this protein to achieve morphologically normal plants that exhibit high levels of Cry1A.105 protein expression and accumulation, and that exhibit commercial levels of resistance to a broad range of Lepidopteran insect plant pests in the genus' selected from the groups consisting of *Anticarsia*, *Pseudoplusia*, *Rachiplusia*, *Helicoverpa*, *Heliothis*, *Spodoptera*, *Epinotia*, and *Armigera*. It is believed that any plastid targeting peptide coding sequence would function effectively for directing the precursor Cry1A.105 protein to the plastid/chloroplast.

Untranslated leader sequences, introns and 3' transcription termination and polyadenylation sequences are known in the art, and the skilled artisan would understand that in certain circumstances, expression can be enhanced or stabilized by incorporating these sequences into the expression cassettes. A number of such sequences are known in the art and are intended to be included within the scope of the present disclosure. Similarly, promoters that function to achieve the regulated expression of a linked sequence are known in

the art and are also intended to be included within the scope of the present disclosure. Promoters can be selected for use to drive expression of a linked sequence in any number of combinations of parameters, including but not limited to temporal control of expression, spatial or tissue specific control of expression, and to control the amount of a particular gene product desired to be accumulated within a particular plant cell or tissue.

The isolated and purified protein comprising an insecticidal fragment of the Cry1A.105 amino acid sequence is also intended to be within the scope of the present invention. Variants are also intended to be within the scope of the invention so long as the amino acid substitution or substitutions effecting the variation are generally conservative with respect to the substituted amino acid(s), and the substitution(s) does not result in a reduction of insecticidal bioactivity or range of species specificity. It is intended that an insecticidal fragment of a Cry1A.105 protein is a segment of the amino acid sequence as set forth in SEQ ID NO:2 from about amino acid position 1 through about amino acid position 650, or from about amino acid position 2 through about amino acid position 612, or from about amino acid position 5 through about amino acid position 610, or from about amino acid position 10 through about amino acid position 600. Alternatively, it is intended that an insecticidal fragment of a Cry1A.105 protein consist of from about 550 to about 650 contiguous amino acids selected from the group consisting of amino acid residues 1 through about 650 as set forth at SEQ ID NO:2. The full length precursor protein, consisting of amino acid residue 1 through about residue 3534, exhibits excellent crystal formation properties and is tolerated well by both monocot and dicot plant species. The precursor protein also exhibits excellent stability when in crystalline form, and also exhibits excellent solubility at alkaline pH, in particular alkaline pH within a range of from about 8.0 to about 12.0, or from about 8.5 to about 11.5, or from about pH 9.0 to about pH 11.0.

The protein of the present invention can be purified and used alone in an insecticidally effective amount in any number of compositions intended for use as a lepidopteran pest control agent, or can be combined in an insecticidally effective amount with any number of other pesticidal agents that are different from the Cry1A.105 protein. Such other pesticidal agents are intended to include but not to be limited to other Bt Cry or other insecticidal compositions whether or not toxic to a lepidopteran species including chemical insecticides, fungicidal or fungistatic agents, antibiotics, antibacterial agents, bacteriostatic agents, and nematocidal or nematostatic agents. Such pesticidal combinations including a Cry1A.105 along with any number of other pesticidal agents can be produced by a transgenic cell, or formulated using purified or substantially purified pesticidal agents into a pesticide

composition in a form consisting of a dust, a granular material, an oil suspension, a water suspension, a mixture of oil and water emulsion, or a wettable powder, and then provided in an agriculturally acceptable carrier for foliar applications. The compositions can be formulated into a seed treatment as well, either together with a Cry1A.105 in the composition intended for inclusion in the seed treatment, or as a composition applied to a seed that is derived from a transgenic plant transformed to express insecticidally effective amounts of a Cry1A.105, so that the seed treatment composition containing pesticidal agents is provided to a target lepidopteran pest along with cells of a plant grown from the seed that are producing pesticidally effective amounts of a Cry1A.105 protein. A combination of insecticidal proteins the each are toxic to the same insect species and yet manifest their toxicity effects through different modes of action would be a particularly useful combination of pesticidal agents for controlling lepidopteran species or delaying the onset of resistance to any single pesticidal agent otherwise effective against a particular lepidopteran species. An exemplary combination of such proteins would be a Cry1A.105 protein of the present invention, i.e., a first insecticidal protein, coupled with at least a second insecticidal protein different from the first. Such different insecticidal proteins include but are not limited to other lepidopteran Bt. crystalline proteins (other Cry1's, Cry2's, Cry5's, Cry9's), VIP proteins, lepidopteran insecticidal proteins referred to as TIC proteins, and insecticidal proteins produced by *Xenorhabdus* and *Photorhabdus* species of bacteria. Providing in the diet of an insect pest a combination of one or more insecticidal proteins along with an agent designed for achieving dsRNA mediated gene suppression of one or more genes essential for insect survival is a particularly useful combination of pesticidal agents for controlling lepidopteran species or delaying the onset of resistance to any single pesticidal agent otherwise effective against a particular lepidopteran species.

Plants transformed with the nucleotide sequences of the present invention are provided as another embodiment of the present invention. Methods for stably introducing DNA into plant cells is known in the art, and includes but is not limited to vacuum infiltration, *Agrobacterium* or *Rhizobium* mediated transformation, electroporation, and various ballistic methods. DNA introduced into plants is generally targeted for insertion into the nuclear chromosome, although insertion into the chloroplast or plastid DNA can be achieved. DNA introduced into plants is generally linked to or associated with a sequence that provides a means for identifying or selecting the cell or cells that have been stably transformed with the DNA of interest, including but not limited to scoreable markers such as fluorescence or light emitting genes and genes encoding pigments or enzymes that, in the

presence of the appropriate substrate, impart a colorimetric feature to the transformed cell or cells, or by including selectable markers that allow a positive selection of transformed cells and tissue, providing a growth advantage to the transformed cells and essentially causing the non-transformed cells or tissue to become static or to die. Such selectable markers include but are not limited to genes encoding basta, bar, methotrexate resistance, neomycin phosphotransferase, glyphosate insensitive EPSPS enzymes, glyphosate oxidoreductase (GOX) enzymes, *E. coli* phnO or its equivalent, and the like.

Vectors and other types of sequences designed for maintaining, manipulating, and/or shepherding the exemplified nucleotide sequences while being manipulated in the laboratory or for introduction into a host cell are also included within the scope of the invention, and are intended to include but not be limited to phages, plasmids, bacmids, yacmids, cosmids, and the like.

Transformed plants are also within the scope of the present invention. Plants transformed to contain a nucleotide sequence encoding at least an insecticidal fragment of a Cry1A.105 protein are specifically enabled by the present disclosure. Both monocot and dicot plants are envisioned to be within the scope of the present invention. Monocots are intended to include but not be limited to corn, wheat, oat, rice, sorghum, milo, buckwheat, rye, grass (fescue, timothy, brome, orchard, St. Augustine, Bermuda, bentgrass), and barley, and dicot plants are intended to include at least alfalfa, apple, apricot, asparagus, bean, berry, blackberry, blueberry, canola, carrot, cauliflower, celery, cherry, chickpea, citrus tree, cotton, cowpea, cranberry, cucumber, cucurbit, egg plant, fruit tree, grape, lemon, lettuce, linseed, melon, mustard, nut bearing tree, okra, orange, pea, peach, peanut, pear, plum, potato, soybeans, squash, strawberry, sugar beet, sunflower, sweet potato, tobacco, tomato, turnip, and vegetable. Produce from these plants as well as seeds and tissues produced from these plants are specifically included within the present invention, so long as the seed, tissue, or produce contains a transgene encoding an insecticidal fragment of a Cry1A.105 protein.

Methods for detecting, in a biological sample, a Cry1A.105 protein or a nucleotide sequence encoding an insecticidal fragment of a Cry1A.105 protein are provided by the present invention. Cry1A.105 can be used to immunize animals to produce antibodies specific for Cry1A.105 epitopes. Cry1A.105 specific antibodies can be used to detect the presence of Cry1A.105 in a biological sample. Methods for detecting the binding of an antibody to an antigen are known in the art. Detecting the binding of an antibody to a Cry1A.105 epitope in a biological sample is diagnostic for the presence of the protein in the sample.

Nucleotide sequences encoding a Cry1A.105 insecticidal fragment can be detected as well. Synthetic nucleotide probes can be used to bind to a target sequence, i.e., a nucleotide sequence encoding a Cry1A.105 insecticidal fragment. Methods for detecting the binding of a probe to a target sequence are known in the art. Detecting the binding of a probe to the target Cry1A.105 coding sequence is diagnostic for the presence of the coding sequence in the sample.

Synthetic nucleotide primers can be used in thermal amplification reactions to produce an amplicon from a biological sample suspected of containing a nucleotide sequence encoding an insecticidal fragment of a Cry1A.105 protein. The presence of an amplicon produced in such a thermal amplification reaction is diagnostic for the presence of the nucleotide sequence in the sample. Particularly useful sequences as probes which are diagnostic for detecting the presence of the Cry1A.105 coding sequences of the present invention in a biological sample are sequences that correspond to or are perfectly complementary to (1) nucleotide position 1401-1420 as set forth at SEQ ID NO:1 or SEQ ID NO:3, or (2) nucleotide position 1821 – 1840 as set forth at SEQ ID NO:1 or SEQ ID NO:3. These sequences correspond to (1) the 20 nucleotides spanning the sequence encoding the junction between Domain II and Domain III of the segments of different insecticidal proteins used for constructing the insecticidal portion of the proteins of the present invention, and (2) the 20 nucleotides spanning the sequence encoding the junction between Domain III and the protoxin coding segment of the different protein coding segments used for constructing the coding sequence of the pre-pro-toxin Cry1Ab.105 protein. Nucleotide sequences that are, or are complementary to, either of these segments of DNA (1401-1420 or 1821-1840) can be used as probes for detecting the presence of these coding sequences in biological samples. The detecting of such binding is diagnostic for the presence of such coding sequences in a biological sample. Other sequences as will be recognized by the skilled artisan that flank either side of these segments of DNA can be used as primers for amplifying various sized amplicon segments from such biological samples, and such amplicons are diagnostic for the presence of such coding sequences in the sample. For example, a first primer sequence corresponding to the nucleotide sequence set forth at SEQ ID NO:1 from position 1201-1220 could be used as a forward primer in a thermal amplification reaction with a second primer sequence corresponding to the reverse complement of the nucleotide sequence as set forth at SEQ ID NO:1 from position 1581-1600. Such primers when used together in a thermal amplification reaction with a biological sample containing SEQ ID NO:1 would result in the synthesis of an amplicon corresponding to SEQ ID NO:1 from nucleotide position 1201

through 1600, i.e., a 400 nucleotide amplicon, which would contain the 20 nucleotide segment from nucleotide position 1401 – 1420 as set forth in SEQ ID NO:1, and would therefore be diagnostic for the presence of the Cry1A.105 coding sequence in such sample.

5 A kit for detecting the presence of a Cry1A.105 or detecting the presence of a nucleotide sequence encoding a Cry1A.105 in a sample is provided. The kit is provided along with all reagents and control samples necessary for carrying out a method for detecting the intended agent, as well as instructions for use.

10 The following examples describe preferred embodiments of the invention. Other embodiments within the scope of the claims will be apparent to one skilled in the art from consideration of the specification or practice of the invention as disclosed herein. It is intended that the specification, together with the examples, be considered exemplary only, with the scope and spirit of the invention being indicated by the claims which follow the examples.

EXAMPLES

Example 1.

This example illustrates synthetic nucleotide sequences encoding an insecticidal
5 Cry1A.105 protein.

A nucleotide sequence as set forth at SEQ ID NO:1 encoding a Cry1A.105 insecticidal protein was constructed for use in dicotyledonous plants. The amino acid sequence translation is set forth at SEQ ID NO:2. The toxin encoding segment consists of nucleotides from about position 1 through about position 1830, more or less.

10 A nucleotide sequence as set forth at SEQ ID NO:3 encoding a Cry1A.105 amino acid sequence was constructed for expression in monocotyledonous plants. The amino acid sequence translation is set forth at SEQ ID NO:4. The toxin encoding segment consists of nucleotide from about position 1 through about position 1830, more or less.

The nucleotide sequences as set forth at SEQ ID NO:1 and SEQ ID NO:3 are
15 substantial equivalents of each other. SEQ ID NO:1 and SEQ ID NO:3 exhibit an overall identity of about 94.3%. The two coding sequences are identical from about nucleotide position 1330 through the nucleotide position 3534. The toxin encoding portion of each sequence consists of from about nucleotide position 1 through nucleotide position 1830, and these segments exhibit about 88.9% identity to each other. The substantial differences
20 between the two sequences lie within from about nucleotide position 1 through about nucleotide position 1329, or about the first two thirds of the segment encoding the toxin portion of the Cry1A.105 protein. The two sequences exhibit about 84.7% identity throughout this segment.

An *E. coli* strain (TOP10, Invitrogen, Inc.) transformed with a plasmid designated as
25 pMON70522 containing a beta-lactamase selectable marker and a sequence as set forth at SEQ ID NO:3 encoding a Cry1A.105 was deposited on August 31, 2005, with the Agriculture Research Culture Collection (NRRL) International Depository Authority at 1815 North University Street, in Peoria, Illinois 61604 U.S.A., according to the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent
30 Procedures and was designated as NRRL B-30873.

Example 2.

This example illustrates transgenic cotton plants expressing a Cry1A.105 protein.

Delta and Pineland DP50 cotton seeds were surface sterilized and germinated overnight. Meristem explants were isolated and the primary leaves were removed by micro dissection. Dissected explants were placed in a targeting medium such that the meristems were oriented perpendicular to the direction of the particle delivery. The transformation vector, pMON47740, comprises an expression cassette having a nucleotide sequence set forth in SEQ ID NO:9. A *KpnI* fragment containing a GUS marker gene under the control of an e35S promoter and a chloroplast targeted Cry1A.105 coding sequence under the control of an eFMV promoter was excised from this plasmid and isolated by HPLC and used for gun transformation of the cotton meristem explants. Purified DNA containing both the Cry1A.105 expression cassette and the GUS marker was precipitated onto microscopic gold beads and coated in a thin layer onto a Mylar sheet. The DNA was accelerated into the meristem tissue by electric discharge particle delivery under a partial vacuum. Following bombardment, explants were de-targeted onto hormone-free media without a selective agent. Leaf tissues from regenerated plantlets were sampled and assayed for expression of the GUS marker. Transgenic plants exhibiting a high level of GUS expression were sent to the greenhouse for further screens. These plants were again tested for expression of GUS and negative portions of the plants were pruned off. This cycle of sampling and pruning of GUS-negative tissues was repeated until all sectors of from each plant were positive for the GUS marker. The plants were then maintained under standard greenhouse conditions until seed harvest.

Tissues obtained from F1 GUS positive transgenic cotton plants were tested in bioassays for insecticidal activity against cotton bollworm (CBW) and fall armyworm (FAW). Previously generated isogenic cotton plants expressing insecticidal levels of Cry1Ac or a combination of Cry1Ac and Cry2Ab were used as positive controls and a non-transgenic isoline was used as the negative control.

CBW square assays were used as one means for determining insecticidal activity of the transgenic cotton plants. (Adamczyk et al., (2001) J. Econ. Entomol. 94:284-290; Kranthi et al (2005) Current Science 89:291-298). Squares of leaf tissue (match head size or larger) were collected and placed individually in assay wells. Each square was infested with a single third-instar CBW larva. The number of surviving insects was recorded five days after infestation.

CBW boll assays were also used to determine the insecticidal activity of boll tissue collected from the transgenic plants. 8 hard green bolls (post bloom) from each event were

collected and placed in individual cups and infested with third instar CBW larvae. The number of surviving insects was recorded five days after infestation.

Leaf assays were conducted to determine the insecticidal activity of transgenic leaf tissue against FAW. New leaves were taken from terminals of cotton plants. 2 leaf punches, each about 3/4" in diameter, were collected and placed in each of 16 individual assay wells. Each well was infested with a single second or third instar FAW larva. The number of surviving insects was recorded five days after infestation.

Bioassay results are shown in Table 1. The results show that transgenic cotton events expressing Cry1A.105 exhibited greater insecticidal activity than transgenic events expressing either Cry1Ac or a combination of Cry1Ac and Cry2Ab against both FAW and CBW.

Table 1. Bioassay results of FAW and CBW using the transgenic cotton plant tissue.

Plant	FAW (% survival) (leaf tissue)	CBW (% survival) (Square tissue)	CBW (% survival) (Boll tissue)
Cry1Ac/Cry2Ab	74.5	32.0	35.8
Cry1Ac	92.7	35.5	35.8
Isoline	99.6	96.8	54
17238	10.9	9.4	25
17567	0	12.5	12.5
17774	1.6	1.2	0
17875	3.1	4.2	0
18026	1.6	18.8	12.5
18122	7.8	22.9	0

Tobacco budworm and corn earworms were also tested in similar bioassays. In each case, the Cry1A.105 plants exhibited insecticidal activity against these pests as well.

Example 3.

This example illustrates transgenic corn plants expressing a Cry1A.105 protein.

Transgenic corn plants were regenerated from cells transformed with the vector pMON40232. pMON40232 contains an expression cassette having a nucleotide sequence as set forth in SEQ ID NO:7 that contains, in operable linkage, an enhanced CAMV 35S promoter, a wheat CAB leader sequence, a rice actin 1 intron, a Cry1A.105 coding sequence

and a wheat hsp17 gene 3' transcription termination and polyadenylation sequence. A nucleotide sequence encoding an *Arabidopsis thaliana* EPSPS chloroplast targeting sequence (At.EPSES-CTP2) is positioned upstream of and in frame with the Cry1A.105 coding sequence. pMON40232 contains a recombinant gene encoding an EPSPS that is insensitive to the herbicide glyphosate, for use in selection of transgenic events. Transgenic events arising from tissue transformed with pMON40232 were designated as LAJ 105. Transgenic events were screened for the absence of any vector backbone, for the presence of a single simple inserted sequence, and for the intactness of the expression cassette containing the nucleotide sequence encoding the Cry1A.105 protein.

Bioassays were conducted with events that met the limitations of the event screen. LAJ105 transgenic corn plants were compared in the bioassay to an isogenic LH198 negative control and a positive control MON810 variety expressing the insecticidal portion of a Cry1Ab protein. Five leaf disks, each about one centimeter in diameter, were obtained from each of 10 individual Cry1A.105 transgenic events and from the controls. Leaf disks were placed on agar filled wells to keep the plant material turgid. Discs were then subjected to feeding by FAW, black cutworm (BCW), European corn borer (ECB), corn ear worm (CEW), and Southwestern corn borer (SWCB) neonate larvae. A single neonate FAW larvae, a single CEW larvae, two neonate BCW, two neonate SWCB larvae, or four neonate ECB larvae were applied to each well. Feeding damage was evaluated after four days, using a leaf damage rating (LDR) scale from 0 - 11, 0 indicating no visible feeding damage, 11 indicating at least 50% of the disc was eaten, and each point on the scale between 0 and 11 indicating a 5% increase in observed feeding damage to the leaf disc under observation.

Bioassay results indicated that events expressing Cry1A.105 protein exhibited greater insecticidal activity toward FAW, ECB and CEW than the LDR's exhibited by the Cry1Ab control against the same pest larvae. LDR's for these three pests on the Cry1A.105 events was less than 1 while the Cry1Ab control exhibited LDR's ranging from about 8 to about 10. The LDR was consistently between 1 and 2 both for the Cry1A.105 events and for the Cry1Ab control when tested for activity against SWCB, indicating that the Cry1A.105 protein was no more toxic to SWCB than was Cry1Ab. The results of this bioassay supported previous results that indicated that Cry1Ab is ineffective in controlling BCW. The Cry1A.105 events were no more effective against BCW than was the Cry1Ab control. Thus, at the levels of expression of the Cry1A.105 protein *in planta*, these plants would be effective in controlling other lepidopteran genus plant pests including but not

limited to those in the genus *Anticarsia*, *Pseudoplusia*, *Rachiplusia*, *Heliothis*, *Helicoverpa*, *Spodoptera*, *Epinotia*, and *Armigera*.

Claims

What is claimed is:

1. An isolated polynucleotide encoding an insecticidal protein comprising an amino acid sequence as set forth in SEQ ID NO:2 from amino acid position 10 through amino acid position 600.
2. The isolated polynucleotide of Claim 1 selected from the group consisting of SEQ ID NO:1 and SEQ ID NO:3.
3. The isolated polynucleotide of Claim 2 for use in expressing said insecticidal protein in a crop plant.
4. The isolated polynucleotide of Claim 3 wherein said crop plant is selected from the group consisting of a monocotyledonous crop plant and a dicotyledonous crop plant.
5. The isolated polynucleotide of Claim 4 wherein said monocotyledonous plant is selected from the group of plants consisting of corn, wheat, oat, rice, sorghum, milo, buckwheat, rye, grass (fescue, timothy, brome, orchard, St. Augustine, Bermuda, bentgrass), and barley.
6. The isolated polynucleotide of Claim 4 wherein said dicotyledonous plant is selected from the group of plants consisting of alfalfa, apple, apricot, asparagus, bean, berry, blackberry, blueberry, canola, carrot, cauliflower, celery, cherry, chickpea, citrus tree, cotton, cowpea, cranberry, cucumber, cucurbit, egg plant, fruit tree, grape, lemon, lettuce, linseed, melon, mustard, nut bearing tree, okra, orange, pea, peach, peanut, pear, plum, potato, soybeans, squash, strawberry, sugar beet, sunflower, sweet potato, tobacco, tomato, turnip, and vegetable.
7. An isolated and purified insecticidal protein comprising an amino acid sequence as set forth in SEQ ID NO: 2 from amino acid position 10 through amino acid position 600.

8. A composition comprising an insecticidally effective amount of the isolated and purified protein of Claim 7.
9. An expression cassette for use in expressing an insecticidal protein having an amino acid sequence as set forth in SEQ ID NO:2 in a host cell, wherein said expression cassette comprises in operable linkage a promoter sequence functional in said host cell and a polynucleotide encoding said protein.
10. The expression cassette of Claim 9 wherein said host cell is selected from the group consisting of a bacterial cell, a fungal cell, a mammalian cell, and a plant cell.
11. The expression cassette of Claim 10 wherein
 - (a) said bacterial cell is selected from the group consisting of a *Bacillus* species cell, a *Enterobacteriaceae* species cell, a *Pseudomonas* species cell, a *Clostridium* species cell, and a *Rhizobium* species cell, and a *Agrobacterium* species cell; and
 - (b) said plant cell is selected from the group of plants consisting of a dicotyledonous plant and a monocotyledonous plant, said dicotyledonous plant being further selected from the group consisting of alfalfa, apple, apricot, asparagus, bean, berry, blackberry, blueberry, canola, carrot, cauliflower, celery, cherry, chickpea, citrus tree, cotton, cowpea, cranberry, cucumber, cucurbit, egg plant, fruit tree, grape, lemon, lettuce, linseed, melon, mustard, nut bearing tree, okra, orange, pea, peach, peanut, pear, plum, potato, soybeans, squash, strawberry, sugar beet, sunflower, sweet potato, tobacco, tomato, turnip, and vegetable, and said monocotyledonous plant being further selected from the group consisting of corn, wheat, oat, rice, sorghum, milo, buckwheat, rye, grass (fescue, timothy, brome, orchard, St. Augustine, Bermuda, bentgrass), and barley.
12. The expression cassette of Claim 9 wherein said host cell is a plant cell and said expression cassette further comprises in operable linkage a polynucleotide selected from the group consisting of an expression enhancer sequence, an untranslated leader sequence, an intron sequence, a chloroplast targeting peptide encoding sequence, and a transcription termination and polyadenylation sequence.
13. A vector comprising the expression cassette of any of Claims 9 through Claim 12.

14. A transgenic plant or plant cell resistant to lepidopteran insect infestation comprising a polynucleotide encoding a protein exhibiting an amino acid sequence as set forth in SEQ ID NO:2 from amino acid position 10 through amino acid position 600, wherein said protein exhibits insecticidal activity when expressed in said plant or plant cell.
15. The transgenic plant or plant cell of claim 14 wherein said transgenic plant is selected from the group consisting of a dicotyledonous plant and a monocotyledonous plant, said dicotyledonous plant being further selected from the group consisting of alfalfa, apple, apricot, asparagus, bean, berry, blackberry, blueberry, canola, carrot, cauliflower, celery, cherry, chickpea, citrus tree, cotton, cowpea, cranberry, cucumber, cucurbit, egg plant, fruit tree, grape, lemon, lettuce, linseed, melon, mustard, nut bearing tree, okra, orange, pea, peach, peanut, pear, plum, potato, soybeans, squash, strawberry, sugar beet, sunflower, sweet potato, tobacco, tomato, turnip, and vegetable, and said monocotyledonous plant being further selected from the group consisting of corn, wheat, oat, rice, sorghum, milo, buckwheat, rye, grass (fescue, timothy, brome, orchard, St. Augustine, Bermuda, bentgrass), and barley.
16. The transgenic plant or plant cell of claim 15 resistant to lepidopteran insect infestation, wherein said lepidopteran insect is selected from the group consisting of a leaf roller, a cutworm, an armyworm, a borer, a bagworm, and a forage feeder.
17. The transgenic plant or plant cell of claim 16, wherein said lepidopteran insect is further selected from the group consisting of a fall armyworm, a European corn borer, a corn earworm or cotton bollworm, a southwestern corn borer, and a black cutworm.
18. Progeny or seed of the transgenic plant or plant cell of claim 15, wherein said progeny or seed comprise said polynucleotide.
19. A method for controlling lepidopteran insect infestation of a plant by providing in the diet of the insect one or more plant cells transformed with a nucleic acid comprising in operable linkage a plant functional promoter and a polynucleotide encoding a protein that exhibits an amino acid sequence as set forth in SEQ ID NO:2 from amino acid position 10 through amino acid position 600 exhibiting lepidopteran insecticidal activity.

20. The expression cassette of claim 12 comprising a polynucleotide selected from the group consisting of SEQ ID NO:5 and SEQ ID NO:7.
21. An isolated polynucleotide encoding an insecticidal protein exhibiting an amino acid sequence as set forth at SEQ ID NO:2 from about amino acid 1 through about amino acid 612.
22. An isolated polynucleotide encoding an insecticidal protein as set forth at SEQ ID NO:2 from about amino acid 1 through about amino acid 610.
23. The isolated polynucleotide of Claim 21 exhibiting at least about 90% identity to the nucleotide sequence as set forth at SEQ ID NO:1.
24. The isolated polynucleotide of Claim 21 exhibiting at least about 90% identity to the nucleotide sequence as set forth at SEQ ID NO:3.
25. A hybrid insecticidal protein comprising an amino acid segment comprising from about 500 to about 600 contiguous amino acids selected from the amino acid segment as set forth at SEQ ID NO:2 from about amino acid position 10 through about amino acid position 600.
26. A composition comprising an insecticidally effective amount of a CryIA.105 protein exhibiting an amino acid sequence as set forth in SEQ ID NO:2 from about amino acid position 10 through about amino acid position 600, wherein said composition is selected from the group consisting of a plant cell, a bacterial cell, a fungal cell, a colloid, an emulsion, a seed coating, a bait, and a powder.
27. The composition of claim 26 wherein said CryIA.105 protein is present in an amount from about 0.5 to about 200 parts per million (PPM).
28. The composition of claim 27 wherein said CryIA.105 protein is present in an amount from about 0.5 to about 20 PPM.

29. The composition of claim 27 wherein said composition is a plant cell or a group of plant cells.
30. A method for protecting a crop in a field from lepidopteran insect infestation comprising growing a transgenic crop plant comprising an insecticidally effective amount of a Cry1A.105 protein insecticidal agent in the diet of the lepidopteran insect to inhibit the insect from surviving on said transgenic crop plant, wherein said insecticidal agent comprises an amino acid sequence from about amino acid position 10 through about amino acid position 600 as set forth in SEQ ID NO:2.
31. The method of claim 30, wherein said transgenic crop plant further comprises an additional insecticidal agent toxic to the same insect species as the Cry1A.105 and is selected from the group consisting of a Bacillus toxin, a Xenorhabdus toxin, a Photorhabdus toxin, and a dsRNA specific for suppression of one or more essential genes in said insect species.
32. The method of claim 30 or 31 wherein the yield of said crop is improved compared to the yield of an isogenic crop lacking said insecticidal agent or agents.

SEQUENCE LISTING

<110> Monsanto Technology LLC
 Bogdanova, Natalia N.
 Corbin, David R.
 Malvar, Thomas M.
 Perlak, Frederick J.
 Roberts, James K.
 Romano, Charles P.

<120> NUCLEOTIDE SEQUENCES ENCODING INSECTICIDAL PROTEINS

<130> 38-21 (52905) A

<150> 60/713,144
 <151> 2005-08-31

<160> 8

<170> PatentIn version 3.3

<210> 1
 <211> 3534
 <212> DNA
 <213> artificial

<220>
 <223> synthetic nucleotide sequence encoding CryIA.105 amino acid sequence

<220>
 <221> CDS
 <222> (1)..(3534)

<400> 1
 atg gac aac aac cca aac atc aac gaa tgc att cca tac aac tgc ttg 48
 Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
 1 5 10 15
 agt aac cca gaa gtt gaa gta ctt ggt gga gaa cgc att gaa acc ggt 96
 Ser Asn Pro Glu Val Glu Val Leu Gly Glu Arg Ile Glu Thr Gly
 20 25 30
 tac act ccc atc gac atc tcc ttg tcc ttg aca cag ttt ctg ctc agc 144
 Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
 35 40 45
 gag ttc gtg cca ggt gct ggg ttc gtt ctc gga cta gtt gac atc atc 192
 Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
 50 55 60
 tgg ggt atc ttt ggt cca tct caa tgg gat gca ttc ctg gtg caa att 240
 Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
 65 70 75 80
 gag cag ttg atc aac cag agg atc gaa gag ttc gcc agg aac cag gcc 288
 Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
 85 90 95
 atc tct agg ttg gaa gga ttg agc aat ctc tac caa atc tat gca gag 336
 Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu
 100 105 110
 agc ttc aga gag tgg gaa gcc gat cct act aac cca gct ctc cgc gag 384
 Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu
 115 120 125
 gaa atg cgt att caa ttc aac gac atg aac agc gcc ttg acc aca gct 432
 Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala
 130 135 140
 atc cca ttg ttc gca gtc cag aac tac caa gtt cct ctc ttg tcc gtg 480
 Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val
 145 150 155 160
 tac gtt caa gca gct aat ctt cac ctc agc gtg ctt cga gac gtt agc 528
 Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser
 165 170 175

gtg ttt ggg caa agg tgg gga ttc gat gct gca acc atc aat agc cgt Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg 180 185 190	576
tac aac gac ctt act agg ctg att gga aac tac acc gac cac gct gtt Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp His Ala Val 195 200 205	624
cgt tgg tac aac act ggc ttg gag cgt gtc tgg ggt cct gat tct aga Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg 210 215 220	672
gat tgg att aga tac aac cag ttc agg aga gaa ttg acc ctc aca gtt Asp Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val 225 230 235 240	720
ttg gac att gtg tct ctc ttc ccg aac tat gac tcc aga acc tac cct Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr Tyr Pro 245 250 255	768
atc cgt aca gtg tcc caa ctt acc aga gaa atc tat act aac cca gtt Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val 260 265 270	816
ctt gag aac ttc gac ggt agc ttc cgt ggt tct gcc caa ggt atc gaa Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu 275 280 285	864
ggc tcc atc agg agc cca cac ttg atg gac atc ttg aac agc ata act Gly Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr 290 295 300	912
atc tac acc gat gct cac aga gga gag tat tac tgg tct gga cac cag Ile Tyr Thr Asp Ala His Arg Gly Glu Tyr Tyr Trp Ser Gly His Gln 305 310 315 320	960
atc atg gcc tct cca gtt gga ttc agc ggg ccc gag ttt acc ttt cct Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro 325 330 335	1008
ctc tat gga act atg gga aac gcc gct cca caa caa cgt atc gtt gct Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala 340 345 350	1056
caa cta ggt cag ggt gtc tac aga acc ttg tct tcc acc ttg tac aga Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg 355 360 365	1104
aga ccc ttc aat atc ggt atc aac aac cag caa ctt tcc gtt ctt gac Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp 370 375 380	1152
gga aca gag ttc gcc tat gga acc tct tct aac ttg cca tcc gct gtt Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val 385 390 395 400	1200
tac aga aag agc gga acc gtt gat tcc ttg gac gaa atc cca cca cag Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln 405 410 415	1248
aac aac aat gtg cca ccc agg caa gga ttc tcc cac agg ttg agc cac Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His 420 425 430	1296
gtg tcc atg ttc cgt tcc gga ttc agc aac agt tcc gtg agc atc atc Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile 435 440 445	1344
aga gct cct atg ttc tct tgg ata cac cgt agt gct gag ttc aac aac Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn 450 455 460	1392
atc att gca tcc gac agc att act caa ata ccc ttg gtg aaa gca cat Ile Ile Ala Ser Asp Ser Ile Thr Gln Ile Pro Leu Val Lys Ala His 465 470 475 480	1440
aca ctt cag tca ggt act act gtt gtc aga ggt cca ggg ttt aca gga Thr Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro Gly Phe Thr Gly 485 490 495	1488

gga gac att ctt cgt cgc aca agt gga gga ccc ttt gct tac act att Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr Ile 500 505 510	1536
gtt aac atc aat ggc caa ttg ccc caa agg tat cgt gca aga atc cgc Val Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg Ala Arg Ile Arg 515 520 525	1584
tat gcc tct act aca aat ctc agg atc tac gtg act gtt gca ggt gaa Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu 530 535 540	1632
agg atc ttt gct ggt cag ttc aac aag act atg gat acc ggt gac cct Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly Asp Pro 545 550 555 560	1680
ttg aca ttc caa tct ttt agc tac gca act atc aac aca gct ttt aca Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr 565 570 575	1728
ttc cca atg agc cag agt agc ttc aca gta ggt gct gac act ttc agc Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe Ser 580 585 590	1776
tca ggg aat gaa gtt tac atc gac agg ttt gaa ttg att cca gtt act Ser Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu Ile Pro Val Thr 595 600 605	1824
gca acc ctc gag gct gag tac aac ctt gag aga gcc cag aag gct gtg Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala Val 610 615 620	1872
aac gcc ctc ttt acc tcc acc aat cag ctt ggc ttg aaa act aac gtt Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val 625 630 635 640	1920
act gac tat cac att gac caa gtg tcc aac ttg gtc acc tac ctt agc Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser 645 650 655	1968
gat gag ttc tgc ctc gac gag aag cgt gaa ctc tcc gag aaa gtt aaa Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys 660 665 670	2016
cac gcc aag cgt ctc agc gac gag agg aat ctc ttg caa gac tcc aac His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn 675 680 685	2064
ttc aaa gac atc aac agg cag cca gaa cgt ggt tgg ggt gga agc acc Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr 690 695 700	2112
ggg atc acc atc caa gga ggc gac gat gtg ttc aag gag aac tac gtc Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val 705 710 715 720	2160
acc ctc tcc gga act ttc gac gag tgc tac cct acc tac ttg tac cag Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln 725 730 735	2208
aag atc gat gag tcc aaa ctc aaa gcc ttc acc agg tat caa ctt aga Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg 740 745 750	2256
ggc tac atc gaa gac agc caa gac ctt gaa atc tac tcg atc agg tac Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Ser Ile Arg Tyr 755 760 765	2304
aat gcc aag cac gag acc gtg aat gtc cca ggt act ggt tcc ctc tgg Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp 770 775 780	2352
cca ctt tct gcc caa tct ccc att ggg aag tgt gga gag cct aac aga Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg 785 790 795 800	2400
tgc gct cca cac ctt gag tgg aat cct gac ttg gac tgc tcc tgc agg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg 805 810 815	2448

gat ggc gag aag tgt gcc cac cat tct cat cac ttc tcc ttg gac atc Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile 820 825 830	2496
gat gtg gga tgt act gac ctg aat gag gac ctc gga gtc tgg gtc atc Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile 835 840 845	2544
ttc aag atc aag acc caa gac gga cac gca aga ctt ggc aac ctt gag Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu 850 855 860	2592
ttt ctc gaa gag aaa cca ttg gtc ggt gaa gct ctc gct cgt gtg aag Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys 865 870 875 880	2640
aga gca gag aag aag tgg agg gac aaa cgt gag aaa ctc gaa tgg gaa Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu 885 890 895	2688
act aac atc gtt tac aag gag gcc aaa gag tcc gtg gat gct ttg ttc Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe 900 905 910	2736
gtg aac tcc caa tat gat cag ttg caa gcc gac acc aac atc gcc atg Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met 915 920 925	2784
atc cac gcc gca gac aaa cgt gtg cac agc att cgt gag gct tac ttg Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu 930 935 940	2832
cct gag ttg tcc gtg atc cct ggt gtg aac gct gcc atc ttc gag gaa Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu 945 950 955 960	2880
ctt gag gga cgt atc ttt acc gca ttc tcc ttg tac gat gcc aga aac Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn 965 970 975	2928
gtc atc aag aac ggt gac ttc aac aat ggc ctc agc tgc tgg aat gtg Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val 980 985 990	2976
aaa ggt cat gtg gac gtg gag gaa cag aac aat cag cgt tcc gtc ctg Lys Gly His Val Asp Val Glu Gln Asn Asn Gln Arg Ser Val Leu 995 1000 1005	3024
gtt gtg cct gag tgg gaa gct gaa gtg tcc caa gag gtt aga gtc Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val 1010 1015 1020	3069
tgt cca ggt aga ggc tac att ctc cgt gtg acc gct tac aag gag Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu 1025 1030 1035	3114
gga tac ggt gag ggt tgc gtg acc atc cac gag atc gag aac aac Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn 1040 1045 1050	3159
acc gac gag ctt aag ttc tcc aac tgc gtc gag gaa gaa atc tat Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr 1055 1060 1065	3204
ccc aac aac acc gtt act tgc aac gac tac act gtg aat cag gaa Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu 1070 1075 1080	3249
gag tac gga ggt gcc tac act agc cgt aac aga ggt tac aac gaa Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu 1085 1090 1095	3294
gct cct tcc gtt cct gct gac tat gcc tcc gtg tac gag gag aaa Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys 1100 1105 1110	3339
tcc tac aca gat ggc aga cgt gag aac cct tgc gag ttc aac aga Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg 1115 1120 1125	3384

ggt tac agg gac tac aca cca ctt cca gtt ggc tat gtt acc aag 3429
 Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys
 1130 1135 1140

gag ctt gag tac ttt cct gag acc gac aaa gtg tgg atc gag atc 3474
 Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile
 1145 1150 1155

ggt gaa acc gag gga acc ttc atc gtg gac agc gtg gag ctt ctc 3519
 Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu
 1160 1165 1170

ttg atg gag gaa taa 3534
 Leu Met Glu Glu
 1175

<210> 2
 <211> 1177
 <212> PRT
 <213> artificial

<220>
 <223> Synthetic Construct

<400> 2

Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
 1 5 10 15

Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly
 20 25 30

Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
 35 40 45

Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
 50 55 60

Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
 65 70 75 80

Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
 85 90 95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu
 100 105 110

Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu
 115 120 125

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala
 130 135 140

Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val
 145 150 155 160

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser
 165 170 175

Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg
 180 185 190

Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp His Ala Val
 195 200 205

Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg

```

210                215                220

Asp Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val
225                230                235                240

Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr Tyr Pro
                245                250                255

Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val
                260                265                270

Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu
                275                280                285

Gly Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr
                290                295                300

Ile Tyr Thr Asp Ala His Arg Gly Glu Tyr Tyr Trp Ser Gly His Gln
305                310                315                320

Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro
                325                330                335

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
                340                345                350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
                355                360                365

Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp
                370                375                380

Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val
385                390                395                400

Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln
                405                410                415

Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His
                420                425                430

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
                435                440                445

Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn
                450                455                460

Ile Ile Ala Ser Asp Ser Ile Thr Gln Ile Pro Leu Val Lys Ala His
465                470                475                480

Thr Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro Gly Phe Thr Gly
                485                490                495

Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr Ile
                500                505                510

Val Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg Ala Arg Ile Arg
                515                520                525

Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu

```

530	535	540
Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly Asp Pro 545 550 555 560		
Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr 565 570 575		
Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe Ser 580 585 590		
Ser Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu Ile Pro Val Thr 595 600 605		
Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala Val 610 615 620		
Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val 625 630 635 640		
Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser 645 650 655		
Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys 660 665 670		
His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn 675 680 685		
Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr 690 695 700		
Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val 705 710 715 720		
Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln 725 730 735		
Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg 740 745 750		
Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Ser Ile Arg Tyr 755 760 765		
Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp 770 775 780		
Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg 785 790 795 800		
Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg 805 810 815		
Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile 820 825 830		
Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile 835 840 845		
Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu		

850 855 860
 Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys
 865 870 875 880
 Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu
 885 890 895
 Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe
 900 905 910
 Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met
 915 920 925
 Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu
 930 935 940
 Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu
 945 950 955 960
 Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn
 965 970 975
 Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val
 980 985 990
 Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu
 995 1000 1005
 Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val
 1010 1015 1020
 Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu
 1025 1030 1035
 Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn
 1040 1045 1050
 Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr
 1055 1060 1065
 Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu
 1070 1075 1080
 Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu
 1085 1090 1095
 Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys
 1100 1105 1110
 Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg
 1115 1120 1125
 Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys
 1130 1135 1140
 Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile
 1145 1150 1155
 Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu

1160 1165 1170

Leu Met Glu Glu
1175

<210> 3
<211> 3534
<212> DNA
<213> artificial

<220>
<223> synthetic nucleotide sequence encoding Cry1A.105 amino acid sequence

<220>
<221> CDS
<222> (1)..(3534)

<400> 3
atg gac aac aac cca aac atc aac gag tgc atc ccg tac aac tgc ctc 48
Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
1 5 10 15
agc aac cct gag gtc gag gtg ctc ggc ggt gag cgc atc gag acc ggt 96
Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly
20 25 30
tac acc ccc atc gac atc tcc ctc tcc ctc acg cag ttc ctg ctc agc 144
Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
35 40 45
gag ttc gtg cca ggc gct ggc ttc gtc ctg ggc ctc gtg gac atc atc 192
Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
50 55 60
tgg ggc atc ttt ggc ccc tcc cag tgg gac gcc ttc ctg gtg caa atc 240
Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
65 70 75 80
gag cag ctc atc aac cag agg atc gag gag ttc gcc agg aac cag gcc 288
Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
85 90 95
atc agc cgc ctg gag ggc ctc agc aac ctc tac caa atc tac gct gag 336
Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu
100 105 110
agc ttc cgc gag tgg gag gcc gac ccc act aac cca gct ctc cgc gag 384
Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu
115 120 125
gag atg cgc atc cag ttc aac gac atg aac agc gcc ctg acc acc gcc 432
Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala
130 135 140
atc cca ctc ttc gcc gtc cag aac tac caa gtc ccg ctc ctg tcc gtg 480
Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val
145 150 155 160
tac gtc cag gcc gcc aac ctg cac ctc agc gtg ctg agg gac gtc agc 528
Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser
165 170 175
gtg ttt ggc cag agg tgg ggc ttc gac gcc gcc acc atc aac agc cgc 576
Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg
180 185 190
tac aac gac ctc acc agg ctg atc ggc aac tac acc gac cac gct gtc 624
Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp His Ala Val
195 200 205
cgc tgg tac aac act ggc ctg gag cgc gtc tgg ggc cct gat tct aga 672
Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg
210 215 220
gac tgg att cgc tac aac cag ttc agg cgc gag ctg acc ctc acc gtc 720
Asp Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val

225	230	235	240	
ctg gac att gtg tcc ctc ttc ccg aac tac gac tcc cgc acc tac ccg Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr Tyr Pro	245	250	255	768
atc cgc acc gtg tcc caa ctg acc cgc gaa atc tac acc aac ccc gtc Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val	260	265	270	816
ctg gag aac ttc gac ggt agc ttc agg ggc agc gcc cag ggc atc gag Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu	275	280	285	864
ggc tcc atc agg agc cca cac ctg atg gac atc ctc aac agc atc act Gly Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr	290	295	300	912
atc tac acc gat gcc cac cgc ggc gag tac tac tgg tcc ggc cac cag Ile Tyr Thr Asp Ala His Arg Gly Glu Tyr Tyr Trp Ser Gly His Gln	305	310	315	960
atc atg gcc tcc ccg gtc ggc ttc agc ggc ccc gag ttt acc ttt cct Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro	325	330	335	1008
ctc tac ggc acg atg ggc aac gcc gct cca caa caa cgc atc gtc gct Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala	340	345	350	1056
cag ctg ggc cag ggc gtc tac cgc acc ctg agc tcc acc ctg tac cgc Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg	355	360	365	1104
agg ccc ttc aac atc ggt atc aac aac cag cag ctg tcc gtc ctg gat Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp	370	375	380	1152
ggc act gag ttc gcc tac ggc acc tcc tcc aac ctg ccc tcc gct gtc Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val	385	390	395	1200
tac cgc aag agc ggc acg gtg gat tcc ctg gac gag atc cca cca cag Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln	405	410	415	1248
aac aac aat gtg ccc ccc agg cag ggt ttt tcc cac agg ctc agc cac Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His	420	425	430	1296
gtg tcc atg ttc cgc tcc ggc ttc agc aac tcg tcc gtg agc atc atc Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile	435	440	445	1344
aga gct cct atg ttc tct tgg ata cac cgt agt gct gag ttc aac aac Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn	450	455	460	1392
atc att gca tcc gac agc att act caa ata ccc ttg gtg aaa gca cat Ile Ile Ala Ser Asp Ser Ile Thr Gln Ile Pro Leu Val Lys Ala His	465	470	475	1440
aca ctt cag tca ggt act act gtt gtc aga ggt cca ggg ttt aca gga Thr Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro Gly Phe Thr Gly	485	490	495	1488
gga gac att ctt cgt cgc aca agt gga gga ccc ttt gct tac act att Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr Ile	500	505	510	1536
gtt aac atc aat ggc caa ttg ccc caa agg tat cgt gca aga atc cgc Val Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg Ala Arg Ile Arg	515	520	525	1584
tat gcc tct act aca aat ctc agg atc tac gtg act gtt gca ggt gaa Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu	530	535	540	1632
agg atc ttt gct ggt cag ttc aac aag act atg gat acc ggt gac cct Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly Asp Pro				1680

545	550	555	560	
ttg aca ttc caa tct ttt agc tac gca act atc aac aca gct ttt aca Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr	565	570	575	1728
ttc cca atg agc cag agt agc ttc aca gta ggt gct gac act ttc agc Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe Ser	580	585	590	1776
tca ggg aat gaa gtt tac atc gac agg ttt gaa ttg att cca gtt act Ser Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu Ile Pro Val Thr	595	600	605	1824
gca acc ctc gag gct gag tac aac ctt gag aga gcc cag aag gct gtg Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala Val	610	615	620	1872
aac gcc ctc ttt acc tcc acc aat cag ctt ggc ttg aaa act aac gtt Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val	625	630	635	1920
act gac tat cac att gac caa gtg tcc aac ttg gtc acc tac ctt agc Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser	645	650	655	1968
gat gag ttc tgc ctc gac gag aag cgt gaa ctc tcc gag aaa gtt aaa Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys	660	665	670	2016
cac gcc aag cgt ctc agc gac gag agg aat ctc ttg caa gac tcc aac His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn	675	680	685	2064
ttc aaa gac atc aac agg cag cca gaa cgt ggt tgg ggt gga agc acc Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr	690	695	700	2112
ggg atc acc atc caa gga ggc gac gat gtg ttc aag gag aac tac gtc Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val	705	710	715	2160
acc ctc tcc gga act ttc gac gag tgc tac cct acc tac ttg tac cag Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln	725	730	735	2208
aag atc gat gag tcc aaa ctc aaa gcc ttc acc agg tat caa ctt aga Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg	740	745	750	2256
ggc tac atc gaa gac agc caa gac ctt gaa atc tac tcg atc agg tac Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Ser Ile Arg Tyr	755	760	765	2304
aat gcc aag cac gag acc gtg aat gtc cca ggt act ggt tcc ctc tgg Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp	770	775	780	2352
cca ctt tct gcc caa tct ccc att ggg aag tgt gga gag cct aac aga Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg	785	790	795	2400
tgc gct cca cac ctt gag tgg aat cct gac ttg gac tgc tcc tgc agg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg	805	810	815	2448
gat ggc gag aag tgt gcc cac cat tct cat cac ttc tcc ttg gac atc Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile	820	825	830	2496
gat gtg gga tgt act gac ctg aat gag gac ctc gga gtc tgg gtc atc Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile	835	840	845	2544
ttc aag atc aag acc caa gac gga cac gca aga ctt ggc aac ctt gag Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu	850	855	860	2592
ttt ctc gaa gag aaa cca ttg gtc ggt gaa gct ctc gct cgt gtg aag Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys				2640

865	870	875	880	
aga gca gag aag aag tgg agg gac aaa cgt gag aaa ctc gaa tgg gaa Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu 885 890 895				2688
act aac atc gtt tac aag gag gcc aaa gag tcc gtg gat gct ttg ttc Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe 900 905 910				2736
gtg aac tcc caa tat gat cag ttg caa gcc gac acc aac atc gcc atg Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met 915 920 925				2784
atc cac gcc gca gac aaa cgt gtg cac agc att cgt gag gct tac ttg Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu 930 935 940				2832
cct gag ttg tcc gtg atc cct ggt gtg aac gct gcc atc ttc gag gaa Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu 945 950 955 960				2880
ctt gag gga cgt atc ttt acc gca ttc tcc ttg tac gat gcc aga aac Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn 965 970 975				2928
gtc atc aag aac ggt gac ttc aac aat ggc ctc agc tgc tgg aat gtg Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val 980 985 990				2976
aaa ggt cat gtg gac gtg gag gaa cag aac aat cag cgt tcc gtc ctg Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu 995 1000 1005				3024
gtt gtg cct gag tgg gaa gct gaa gtg tcc caa gag gtt aga gtc Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val 1010 1015 1020				3069
tgt cca ggt aga ggc tac att ctc cgt gtg acc gct tac aag gag Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu 1025 1030 1035				3114
gga tac ggt gag ggt tgc gtg acc atc cac gag atc gag aac aac Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn 1040 1045 1050				3159
acc gac gag ctt aag ttc tcc aac tgc gtc gag gaa gaa atc tat Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr 1055 1060 1065				3204
ccc aac aac acc gtt act tgc aac gac tac act gtg aat cag gaa Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu 1070 1075 1080				3249
gag tac gga ggt gcc tac act agc cgt aac aga ggt tac aac gaa Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu 1085 1090 1095				3294
gct cct tcc gtt cct gct gac tat gcc tcc gtg tac gag gag aaa Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys 1100 1105 1110				3339
tcc tac aca gat ggc aga cgt gag aac cct tgc gag ttc aac aga Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg 1115 1120 1125				3384
ggt tac agg gac tac aca cca ctt cca gtt ggc tat gtt acc aag Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys 1130 1135 1140				3429
gag ctt gag tac ttt cct gag acc gac aaa gtg tgg atc gag atc Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile 1145 1150 1155				3474
ggt gaa acc gag gga acc ttc atc gtg gac agc gtg gag ctt ctc Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu 1160 1165 1170				3519
ttg atg gag gaa taa Leu Met Glu Glu				3534

1175

<210> 4
 <211> 1177
 <212> PRT
 <213> artificial

<220>
 <223> Synthetic Construct

<400> 4

Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
 1 5 10 15

Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly
 20 25 30

Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
 35 40 45

Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
 50 55 60

Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
 65 70 75 80

Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
 85 90 95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu
 100 105 110

Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu
 115 120 125

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala
 130 135 140

Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val
 145 150 155 160

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser
 165 170 175

Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg
 180 185 190

Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp His Ala Val
 195 200 205

Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg
 210 215 220

Asp Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val
 225 230 235 240

Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr Tyr Pro
 245 250 255

Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val
 260 265 270

Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu
275 280 285

Gly Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr
290 295 300

Ile Tyr Thr Asp Ala His Arg Gly Glu Tyr Tyr Trp Ser Gly His Gln
305 310 315 320

Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro
325 330 335

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
340 345 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
355 360 365

Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp
370 375 380

Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val
385 390 395 400

Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln
405 410 415

Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His
420 425 430

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
435 440 445

Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn
450 455 460

Ile Ile Ala Ser Asp Ser Ile Thr Gln Ile Pro Leu Val Lys Ala His
465 470 475 480

Thr Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro Gly Phe Thr Gly
485 490 495

Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr Ile
500 505 510

Val Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg Ala Arg Ile Arg
515 520 525

Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu
530 535 540

Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly Asp Pro
545 550 555 560

Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr
565 570 575

Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe Ser
580 585 590

Ser Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu Ile Pro Val Thr
 595 600 605
 Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala Val
 610 615 620
 Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val
 625 630 635 640
 Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser
 645 650 655
 Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
 660 665 670
 His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn
 675 680 685
 Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr
 690 695 700
 Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
 705 710 715 720
 Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
 725 730 735
 Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg
 740 745 750
 Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Ser Ile Arg Tyr
 755 760 765
 Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
 770 775 780
 Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg
 785 790 795 800
 Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg
 805 810 815
 Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile
 820 825 830
 Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile
 835 840 845
 Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu
 850 855 860
 Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys
 865 870 875 880
 Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu
 885 890 895
 Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe
 900 905 910

Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met
 915 920 925
 Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu
 930 935 940
 Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu
 945 950 955 960
 Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn
 965 970 975
 Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val
 980 985 990
 Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu
 995 1000 1005
 Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val
 1010 1015 1020
 Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu
 1025 1030 1035
 Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn
 1040 1045 1050
 Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr
 1055 1060 1065
 Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu
 1070 1075 1080
 Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu
 1085 1090 1095
 Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys
 1100 1105 1110
 Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg
 1115 1120 1125
 Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys
 1130 1135 1140
 Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile
 1145 1150 1155
 Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu
 1160 1165 1170
 Leu Met Glu Glu
 1175

<210> 5

<211> 5480

<212> DNA

<213> artificial

<220>

<223> expression cassette encoding CryIA.105 amino acid sequence

<220>
 <221> promoter
 <222> (1)..(981)
 <223> FMV.e35S promoter

 <220>
 <221> 5'UTR
 <222> (985)..(1080)

 <220>
 <221> transit_peptide
 <222> (1090)..(1263)
 <223> chloroplast targeting peptide sequence

 <220>
 <221> CDS
 <222> (1267)..(4797)
 <223> coding sequence for CryIA.105

 <220>
 <221> terminator
 <222> (4833)..(5480)
 <223> Ps.E9 rbcS transcription termination and polyadenylation sequence

 <400> 5
 aattctcagt ccaaagcctc aacaaggta ggtacagag tctccaaacc attagccaaa 60
 agctacagga gatcaatgaa gaatcttcaa tcaaagtaaa ctactgttcc agcacatgca 120
 tcatggtcag taagtttcag aaaaagacat ccaccgaaga cttaaagta gtgggcatct 180
 ttgaaagtaa tcttgtaaac atcgagcagc tggcttggtg ggaccagaca aaaaaggaat 240
 ggtgcagaat tgtaggcgc acctaccaa agcatctttg cctttattgc aaagataaag 300
 cagattcctc tagtacaagt ggggaacaaa ataacgtgga aaagagctgt cctgacagcc 360
 cactcactaa tgcgtatgac gaacgcagtg acgaccacaa aagaattagc ttgagctcag 420
 gatttagcag cattccagat tgggttcaat caacaaggta cgagccatat cactttattc 480
 aaattggtat cgccaaaacc aagaaggaa tcccatcctc aaaggtttgt aaggaagaat 540
 tctcagcca aagcctcaac aaggtcaggg tacagagtct ccaaaccatt agccaaaagc 600
 tacaggagat caatgaagaa tcttcaatca aagtaacta ctgttccagc acatgcatca 660
 tggtcagtaa gtttcagaaa aagacatcca ccgaagactt aaagttagtg ggcattcttg 720
 aaagtaatct tgtcaacatc gagcagctgg ctgtgggga ccagacaaaa aaggaatggt 780
 gcagaattgt taggcgcacc taccaaaaagc atctttgcct ttattgcaa gataaagcag 840
 attcctctag tacaagtggg gaacaaaata acgtggaaaa gagctgtcct gacagccac 900
 tcaactatgc gtatgacgaa cgagtgacg accacaaaag aattccctct atataagaag 960
 gcattcattc ccatttgaag gacacagaaa aatttgctac attgtttcac aaacttcaaa 1020
 tattattcat ttatttgtca gctttcaaac tctttgtttc ttgtttgttg attgagaata 1080
 tttaaaacaa tggcttcctc tatgctctct tccgctacta tgggtgcctc tccggctcag 1140
 gccactatgg tcgctccttt caacggactt aagtcctccg ctgccttccc agccaccgc 1200
 aaggctaaca acgacattac ttccatcaca agcaacggcg gaagagttaa ctgcatgcag 1260
 gccatg gac aac aac cca aac atc aac gaa tgc att cca tac aac tgc 1308
 Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys
 1 5 10
 ttg agt aac cca gaa gtt gaa gta ctt ggt gga gaa cgc att gaa acc 1356
 Leu Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr
 15 20 25 30
 ggt tac act ccc atc gac atc tcc ttg tcc ttg aca cag ttt ctg ctc 1404
 Gly Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu
 35 40 45

agc gag ttc gtg cca ggt gct ggg ttc gtt ctc gga cta gtt gac atc Ser Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile 50 55 60	1452
atc tgg ggt atc ttt ggt cca tct caa tgg gat gca ttc ctg gtg caa Ile Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln 65 70 75	1500
att gag cag ttg atc aac cag agg atc gaa gag ttc gcc agg aac cag Ile Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln 80 85 90	1548
gcc atc tct agg ttg gaa gga ttg agc aat ctc tac caa atc tat gca Ala Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala 95 100 105 110	1596
gag agc ttc aga gag tgg gaa gcc gat cct act aac cca gct ctc cgc Glu Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg 115 120 125	1644
gag gaa atg cgt att caa ttc aac gac atg aac agc gcc ttg acc aca Glu Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr 130 135 140	1692
gct atc cca ttg ttc gca gtc cag aac tac caa gtt cct ctc ttg tcc Ala Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser 145 150 155	1740
gtg tac gtt caa gca gct aat ctt cac ctc agc gtg ctt cga gac gtt Val Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val 160 165 170	1788
agc gtg ttt ggg caa agg tgg gga ttc gat gct gca acc atc aat agc Ser Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser 175 180 185 190	1836
cgt tac aac gac ctt act agg ctg att gga aac tac acc gac cac gct Arg Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp His Ala 195 200 205	1884
gtt cgt tgg tac aac act ggc ttg gag cgt gtc tgg ggt cct gat tct Val Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser 210 215 220	1932
aga gat tgg att aga tac aac cag ttc agg aga gaa ttg acc ctc aca Arg Asp Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr 225 230 235	1980
gtt ttg gac att gtg tct ctc ttc ccg aac tat gac tcc aga acc tac Val Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr Tyr 240 245 250	2028
cct atc cgt aca gtg tcc caa ctt acc aga gaa atc tat act aac cca Pro Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro 255 260 265 270	2076
gtt ctt gag aac ttc gac ggt agc ttc cgt ggt tct gcc caa ggt atc Val Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile 275 280 285	2124
gaa ggc tcc atc agg agc cca cac ttg atg gac atc ttg aac agc ata Glu Gly Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile 290 295 300	2172
act atc tac acc gat gct cac aga gga gag tat tac tgg tct gga cac Thr Ile Tyr Thr Asp Ala His Arg Gly Glu Tyr Tyr Trp Ser Gly His 305 310 315	2220
cag atc atg gcc tct cca gtt gga ttc agc ggg ccc gag ttt acc ttt Gln Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe 320 325 330	2268
cct ctc tat gga act atg gga aac gcc gct cca caa caa cgt atc gtt Pro Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val 335 340 345 350	2316
gct caa cta ggt cag ggt gtc tac aga acc ttg tct tcc acc ttg tac Ala Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr 355 360 365	2364

aga aga ccc ttc aat atc ggt atc aac aac cag caa ctt tcc gtt ctt Arg Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu 370 375 380	2412
gac gga aca gag ttc gcc tat gga acc tct tct aac ttg cca tcc gct Asp Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala 385 390 395	2460
gtt tac aga aag agc gga acc gtt gat tcc ttg gac gaa atc cca cca Val Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro 400 405 410	2508
cag aac aac aat gtg cca ccc agg caa gga ttc tcc cac agg ttg agc Gln Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser 415 420 425 430	2556
cac gtg tcc atg ttc cgt tcc gga ttc agc aac agt tcc gtg agc atc His Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile 435 440 445	2604
atc aga gct cct atg ttc tct tgg ata cac cgt agt gct gag ttc aac Ile Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn 450 455 460	2652
aac atc att gca tcc gac agc att act caa ata ccc ttg gtg aaa gca Asn Ile Ile Ala Ser Asp Ser Ile Thr Gln Ile Pro Leu Val Lys Ala 465 470 475	2700
cat aca ctt cag tca ggt act act gtt gtc aga ggt cca ggg ttt aca His Thr Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro Gly Phe Thr 480 485 490	2748
gga gga gac att ctt cgt cgc aca agt gga gga ccc ttt gct tac act Gly Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr 495 500 505 510	2796
att gtt aac atc aat ggc caa ttg ccc caa agg tat cgt gca aga atc Ile Val Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg Ala Arg Ile 515 520 525	2844
cgc tat gcc tct act aca aat ctc agg atc tac gtg act gtt gca ggt Arg Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly 530 535 540	2892
gaa agg atc ttt gct ggt cag ttc aac aag act atg gat acc ggt gac Glu Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly Asp 545 550 555	2940
cct ttg aca ttc caa tct ttt agc tac gca act atc aac aca gct ttt Pro Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe 560 565 570	2988
aca ttc cca atg agc cag agt agc ttc aca gta ggt gct gac act ttc Thr Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe 575 580 585 590	3036
agc tca ggg aat gaa gtt tac atc gac agg ttt gaa ttg att cca gtt Ser Ser Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu Ile Pro Val 595 600 605	3084
act gca acc ctc gag gct gag tac aac ctt gag aga gcc cag aag gct Thr Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala 610 615 620	3132
gtg aac gcc ctc ttt acc tcc acc aat cag ctt ggc ttg aaa act aac Val Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn 625 630 635	3180
gtt act gac tat cac att gac caa gtg tcc aac ttg gtc acc tac ctt Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu 640 645 650	3228
agc gat gag ttc tgc ctc gac gag aag cgt gaa ctc tcc gag aaa gtt Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val 655 660 665 670	3276
aaa cac gcc aag cgt ctc agc gac gag agg aat ctc ttg caa gac tcc Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser 675 680 685	3324

aac ttc aaa gac atc aac agg cag cca gaa cgt ggt tgg ggt gga agc Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser 690 695 700	3372
acc ggg atc acc atc caa gga ggc gac gat gtg ttc aag gag aac tac Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr 705 710 715	3420
gtc acc ctc tcc gga act ttc gac gag tgc tac cct acc tac ttg tac Val Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr 720 725 730	3468
cag aag atc gat gag tcc aaa ctc aaa gcc ttc acc agg tat caa ctt Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu 735 740 745 750	3516
aga ggc tac atc gaa gac agc caa gac ctt gaa atc tac tcg atc agg Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Ser Ile Arg 755 760 765	3564
tac aat gcc aag cac gag acc gtg aat gtc cca ggt act ggt tcc ctc Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu 770 775 780	3612
tgg cca ctt tct gcc caa tct ccc att ggg aag tgt gga gag cct aac Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn 785 790 795	3660
aga tgc gct cca cac ctt gag tgg aat cct gac ttg gac tgc tcc tgc Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys 800 805 810	3708
agg gat ggc gag aag tgt gcc cac cat tct cat cac ttc tcc ttg gac Arg Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp 815 820 825 830	3756
atc gat gtg gga tgt act gac ctg aat gag gac ctc gga gtc tgg gtc Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val 835 840 845	3804
atc ttc aag atc aag acc caa gac gga cac gca aga ctt ggc aac ctt Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu 850 855 860	3852
gag ttt ctc gaa gag aaa cca ttg gtc ggt gaa gct ctc gct cgt gtg Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val 865 870 875	3900
aag aga gca gag aag aag tgg agg gac aaa cgt gag aaa ctc gaa tgg Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp 880 885 890	3948
gaa act aac atc gtt tac aag gag gcc aaa gag tcc gtg gat gct ttg Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu 895 900 905 910	3996
ttc gtg aac tcc caa tat gat cag ttg caa gcc gac acc aac atc gcc Phe Val Asn Ser Gln Tyr Asp Gln Leu Ala Asp Thr Asn Ile Ala 915 920 925	4044
atg atc cac gcc gca gac aaa cgt gtg cac agc att cgt gag gct tac Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr 930 935 940	4092
ttg cct gag ttg tcc gtg atc cct ggt gtg aac gct gcc atc ttc gag Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu 945 950 955	4140
gaa ctt gag gga cgt atc ttt acc gca ttc tcc ttg tac gat gcc aga Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg 960 965 970	4188
aac gtc atc aag aac ggt gac ttc aac aat ggc ctc agc tgc tgg aat Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn 975 980 985 990	4236
gtg aaa ggt cat gtg gac gtg gag gaa cag aac aat cag cgt tcc gtc Val Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val 995 1000 1005	4284

ctg gtt gtg cct gag tgg gaa gct gaa gtg tcc caa gag gtt aga Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg 1010 1015 1020	4329
gtc tgt cca ggt aga ggc tac att ctc cgt gtg acc gct tac aag Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys 1025 1030 1035	4374
gag gga tac ggt gag ggt tgc gtg acc atc cac gag atc gag aac Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn 1040 1045 1050	4419
aac acc gac gag ctt aag ttc tcc aac tgc gtc gag gaa gaa atc Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile 1055 1060 1065	4464
tat ccc aac aac acc gtt act tgc aac gac tac act gtg aat cag Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln 1070 1075 1080	4509
gaa gag tac gga ggt gcc tac act agc cgt aac aga ggt tac aac Glu Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn 1085 1090 1095	4554
gaa gct cct tcc gtt cct gct gac tat gcc tcc gtg tac gag gag Glu Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu 1100 1105 1110	4599
aaa tcc tac aca gat ggc aga cgt gag aac cct tgc gag ttc aac Lys Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn 1115 1120 1125	4644
aga ggt tac agg gac tac aca cca ctt cca gtt ggc tat gtt acc Arg Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr 1130 1135 1140	4689
aag gag ctt gag tac ttt cct gag acc gac aaa gtg tgg atc gag Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu 1145 1150 1155	4734
atc ggt gaa acc gag gga acc ttc atc gtg gac agc gtg gag ctt Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu 1160 1165 1170	4779
ctc ttg atg gag gaa taa tgagatccac gatatcctgc aggaattggc Leu Leu Met Glu Glu 1175	4827
cggccagctt tcggttcgtat catcggtttc gacaacgttc gtcaagttca atgcatcagt	4887
ttcattgcgc acacaccaga atcctactga gtttgagtat tatggcattg ggaaaactgt	4947
ttttcttgta ccatttggtg tgcttgtaat ttactgtgtt ttttattcgg ttttcgctat	5007
cgaactgtga aatggaaatg gatggagaag agttaatgaa tgatatggtc cttttgttca	5067
ttctcaaatt aatattattt gttttttctc ttatttggtg tgtgttgaat ttgaaattat	5127
aagagatatg caaacatttt gttttgagta aaaatgtgtc aaatcgtggc ctctaattgac	5187
cgaagttaat atgaggagta aaacacttgt agttgtacca ttatgcttat tcactaggca	5247
acaaatata tttcagacct agaaaagctg caaatgttac tgaatacaag tatgtcctct	5307
tgtgttttag acatttatgg actttccttt atgtaatttt ccagaatcct tgtcagattc	5367
taatcattgc tttataatta tagttatact catggatttg tagttgagta tgaaaatatt	5427
ttttaatgca ttttatgact tgccaattga ttgacaacat gcatcaatcg acc	5480
<210> 6	
<211> 1176	
<212> PRT	
<213> artificial	
<220>	
<223> Synthetic Construct	

<400> 6

Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu Ser
1 5 10 15

Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly Tyr
20 25 30

Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser Glu
35 40 45

Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile Trp
50 55 60

Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile Glu
65 70 75 80

Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala Ile
85 90 95

Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu Ser
100 105 110

Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu Glu
115 120 125

Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala Ile
130 135 140

Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val Tyr
145 150 155 160

Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser Val
165 170 175

Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg Tyr
180 185 190

Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp His Ala Val Arg
195 200 205

Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg Asp
210 215 220

Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val Leu
225 230 235 240

Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr Tyr Pro Ile
245 250 255

Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val Leu
260 265 270

Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu Gly
275 280 285

Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr Ile
290 295 300

Tyr Thr Asp Ala His Arg Gly Glu Tyr Tyr Trp Ser Gly His Gln Ile
305 310 315 320

Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro Leu
 325 330 335
 Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala Gln
 340 345 350
 Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg Arg
 355 360 365
 Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp Gly
 370 375 380
 Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val Tyr
 385 390 395 400
 Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln Asn
 405 410 415
 Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His Val
 420 425 430
 Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile Arg
 435 440 445
 Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn Ile
 450 455 460
 Ile Ala Ser Asp Ser Ile Thr Gln Ile Pro Leu Val Lys Ala His Thr
 465 470 475 480
 Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro Gly Phe Thr Gly Gly
 485 490 495
 Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val
 500 505 510
 Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr
 515 520 525
 Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu Arg
 530 535 540
 Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly Asp Pro Leu
 545 550 555 560
 Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr Phe
 565 570 575
 Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe Ser Ser
 580 585 590
 Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu Ile Pro Val Thr Ala
 595 600 605
 Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala Val Asn
 610 615 620
 Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val Thr
 625 630 635 640

Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser Asp
645 650 655

Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys His
660 665 670

Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn Phe
675 680 685

Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr Gly
690 695 700

Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr
705 710 715 720

Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys
725 730 735

Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg Gly
740 745 750

Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Ser Ile Arg Tyr Asn
755 760 765

Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro
770 775 780

Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys
785 790 795 800

Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg Asp
805 810 815

Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp
820 825 830

Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe
835 840 845

Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe
850 855 860

Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg
865 870 875 880

Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr
885 890 895

Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val
900 905 910

Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile
915 920 925

His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro
930 935 940

Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu
945 950 955 960

Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val
 965 970 975
 Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys
 980 985 990
 Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu Val
 995 1000 1005
 Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys
 1010 1015 1020
 Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly
 1025 1030 1035
 Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr
 1040 1045 1050
 Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr Pro
 1055 1060 1065
 Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu Glu
 1070 1075 1080
 Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu Ala
 1085 1090 1095
 Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys Ser
 1100 1105 1110
 Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg Gly
 1115 1120 1125
 Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys Glu
 1130 1135 1140
 Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly
 1145 1150 1155
 Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu
 1160 1165 1170
 Met Glu Glu
 1175

<210> 7
 <211> 4990
 <212> DNA
 <213> artificial
 <220>
 <223> expression cassette encoding Cry1A.105 amino acid sequence
 <220>
 <221> promoter
 <222> (1)..(614)
 <223> CAMV e35s promoter
 <220>
 <221> 5'UTR
 <222> (650)..(710)

Page 26

agg aac cag gcc atc agc cgc ctg gag ggc ctc agc aac ctc tac caa Arg Asn Gln Ala Ile Ser Arg Leu Gly Leu Ser Asn Leu Tyr Gln 95 100 105	1539
atc tac gct gag agc ttc cgc gag tgg gag gcc gac ccc act aac cca Ile Tyr Ala Glu Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro 110 115 120	1587
gct ctc cgc gag gag atg cgc atc cag ttc aac gac atg aac agc gcc Ala Leu Arg Glu Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala 125 130 135 140	1635
ctg acc acc gcc atc cca ctc ttc gcc gtc cag aac tac caa gtc ccg Leu Thr Thr Ala Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro 145 150 155	1683
ctc ctg tcc gtg tac gtc cag gcc gcc aac ctg cac ctc agc gtg ctg Leu Leu Ser Val Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu 160 165 170	1731
agg gac gtc agc gtg ttt ggc cag agg tgg ggc ttc gac gcc gcc acc Arg Asp Val Ser Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr 175 180 185	1779
atc aac agc cgc tac aac gac ctc acc agg ctg atc ggc aac tac acc Ile Asn Ser Arg Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr 190 195 200	1827
gac cac gct gtc cgc tgg tac aac act ggc ctg gag cgc gtc tgg ggc Asp His Ala Val Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly 205 210 215 220	1875
cct gat tct aga gac tgg att cgc tac aac cag ttc agg cgc gag ctg Pro Asp Ser Arg Asp Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu 225 230 235	1923
acc ctc acc gtc ctg gac att gtg tcc ctc ttc ccg aac tac gac tcc Thr Leu Thr Val Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser 240 245 250	1971
cgc acc tac ccg atc cgc acc gtg tcc caa ctg acc cgc gaa atc tac Arg Thr Tyr Pro Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr 255 260 265	2019
acc aac ccc gtc ctg gag aac ttc gac ggt agc ttc agg ggc agc gcc Thr Asn Pro Val Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala 270 275 280	2067
cag ggc atc gag ggc tcc atc agg agc cca cac ctg atg gac atc ctc Gln Gly Ile Glu Gly Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu 285 290 295 300	2115
aac agc atc act atc tac acc gat gcc cac cgc ggc gag tac tac tgg Asn Ser Ile Thr Ile Tyr Thr Asp Ala His Arg Gly Glu Tyr Tyr Trp 305 310 315	2163
tcc ggc cac cag atc atg gcc tcc ccg gtc ggc ttc agc ggc ccc gag Ser Gly His Gln Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu 320 325 330	2211
ttt acc ttt cct ctc tac ggc acg atg ggc aac gcc gct cca caa caa Phe Thr Phe Pro Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln 335 340 345	2259
cgc atc gtc gct cag ctg ggc cag ggc gtc tac cgc acc ctg agc tcc Arg Ile Val Ala Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser 350 355 360	2307
acc ctg tac cgc agg ccc ttc aac atc ggt atc aac aac cag cag ctg Thr Leu Tyr Arg Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu 365 370 375 380	2355
tcc gtc ctg gat ggc act gag ttc gcc tac ggc acc tcc tcc aac ctg Ser Val Leu Asp Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu 385 390 395	2403
ccc tcc gct gtc tac cgc aag agc ggc acg gtg gat tcc ctg gac gag Pro Ser Ala Val Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu 400 405 410	2451

atc cca cca cag aac aac aat gtg ccc ccc agg cag ggt ttt tcc cac Ile Pro Pro Gln Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His 415 420 425	2499
agg ctc agc cac gtg tcc atg ttc cgc tcc ggc ttc agc aac tcg tcc Arg Leu Ser His Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser 430 435 440	2547
gtg agc atc atc aga gct cct atg ttc tct tgg ata cac cgt agt gct Val Ser Ile Ile Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala 445 450 455 460	2595
gag ttc aac aac atc att gca tcc gac agc att act caa ata ccc ttg Glu Phe Asn Asn Ile Ile Ala Ser Asp Ser Ile Thr Gln Ile Pro Leu 465 470 475	2643
gtg aaa gca cat aca ctt cag tca ggt act act gtt gtc aga ggt cca Val Lys Ala His Thr Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro 480 485 490	2691
ggg ttt aca gga gga gac att ctt cgt cgc aca agt gga gga ccc ttt Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe 495 500 505	2739
gct tac act att gtt aac atc aat ggc caa ttg ccc caa agg tat cgt Ala Tyr Thr Ile Val Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg 510 515 520	2787
gca aga atc cgc tat gcc tct act aca aat ctc agg atc tac gtg act Ala Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr 525 530 535 540	2835
gtt gca ggt gaa agg atc ttt gct ggt cag ttc aac aag act atg gat Val Ala Gly Glu Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp 545 550 555	2883
acc ggt gac cct ttg aca ttc caa tct ttt agc tac gca act atc aac Thr Gly Asp Pro Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn 560 565 570	2931
aca gct ttt aca ttc cca atg agc cag agt agc ttc aca gta ggt gct Thr Ala Phe Thr Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala 575 580 585	2979
gac act ttc agc tca ggg aat gaa gtt tac atc gac agg ttt gaa ttg Asp Thr Phe Ser Ser Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu 590 595 600	3027
att cca gtt act gca acc ctc gag gct gag tac aac ctt gag aga gcc Ile Pro Val Thr Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala 605 610 615 620	3075
cag aag gct gtg aac gcc ctc ttt acc tcc acc aat cag ctt ggc ttg Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu 625 630 635	3123
aaa act aac gtt act gac tat cac att gac caa gtg tcc aac ttg gtc Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val 640 645 650	3171
acc tac ctt agc gat gag ttc tgc ctc gac gag aag cgt gaa ctc tcc Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser 655 660 665	3219
gag aaa gtt aaa cac gcc aag cgt ctc agc gac gag agg aat ctc ttg Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu 670 675 680	3267
caa gac tcc aac ttc aaa gac atc aac agg cag cca gaa cgt ggt tgg Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp 685 690 695 700	3315
ggt gga agc acc ggg atc acc atc caa gga ggc gac gat gtg ttc aag Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys 705 710 715	3363
gag aac tac gtc acc ctc tcc gga act ttc gac gag tgc tac cct acc Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr 720 725 730	3411

tac ttg tac cag aag atc gat gag tcc aaa ctc aaa gcc ttc acc agg Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg 735 740 745	3459
tat caa ctt aga ggc tac atc gaa gac agc caa gac ctt gaa atc tac Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr 750 755 760	3507
tcg atc agg tac aat gcc aag cac gag acc gtg aat gtc cca ggt act Ser Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr 765 770 775 780	3555
ggt tcc ctc tgg cca ctt tct gcc caa tct ccc att ggg aag tgt gga Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly 785 790 795	3603
gag cct aac aga tgc gct cca cac ctt gag tgg aat cct gac ttg gac Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp 800 805 810	3651
tgc tcc tgc agg gat ggc gag aag tgt gcc cac cat tct cat cac ttc Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His Phe 815 820 825	3699
tcc ttg gac atc gat gtg gga tgt act gac ctg aat gag gac ctc gga Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly 830 835 840	3747
gtc tgg gtc atc ttc aag atc aag acc caa gac gga cac gca aga ctt Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu 845 850 855 860	3795
ggc aac ctt gag ttt ctc gaa gag aaa cca ttg gtc ggt gaa gct ctc Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu 865 870 875	3843
gct cgt gtg aag aga gca gag aag aag tgg agg gac aaa cgt gag aaa Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys 880 885 890	3891
ctc gaa tgg gaa act aac atc gtt tac aag gag gcc aaa gag tcc gtg Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val 895 900 905	3939
gat gct ttg ttc gtg aac tcc caa tat gat cag ttg caa gcc gac acc Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr 910 915 920	3987
aac atc gcc atg atc cac gcc gca gac aaa cgt gtg cac agc att cgt Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg 925 930 935 940	4035
gag gct tac ttg cct gag ttg tcc gtg atc cct ggt gtg aac gct gcc Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala 945 950 955	4083
atc ttc gag gaa ctt gag gga cgt atc ttt acc gca ttc tcc ttg tac Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr 960 965 970	4131
gat gcc aga aac gtc atc aag aac ggt gac ttc aac aat ggc ctc agc Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser 975 980 985	4179
tgc tgg aat gtg aaa ggt cat gtg gac gtg gag gaa cag aac aat cag Cys Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln 990 995 1000	4227
cgt tcc gtc ctg gtt gtg cct gag tgg gaa gct gaa gtg tcc caa Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln 1005 1010 1015	4272
gag gtt aga gtc tgt cca ggt aga ggc tac att ctc cgt gtg acc Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr 1020 1025 1030	4317
gct tac aag gag gga tac ggt gag ggt tgc gtg acc atc cac gag Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu 1035 1040 1045	4362

atc gag aac aac acc gac gag ctt aag ttc tcc aac tgc gtc gag 4407
 ile Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu
 1050 1055 1060
 gaa gaa atc tat ccc aac aac acc gtt act tgc aac gac tac act 4452
 Glu Glu Ile Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr
 1065 1070 1075
 gtg aat cag gaa gag tac gga ggt gcc tac act agc cgt aac aga 4497
 Val Asn Gln Glu Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg
 1080 1085 1090
 ggt tac aac gaa gct cct tcc gtt cct gct gac tat gcc tcc gtg 4542
 Gly Tyr Asn Glu Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val
 1095 1100 1105
 tac gag gag aaa tcc tac aca gat ggc aga cgt gag aac cct tgc 4587
 Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys
 1110 1115 1120
 gag ttc aac aga ggt tac agg gac tac aca cca ctt cca gtt ggc 4632
 Glu Phe Asn Arg Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly
 1125 1130 1135
 tat gtt acc aag gag ctt gag tac ttt cct gag acc gac aaa gtg 4677
 Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val
 1140 1145 1150
 tgg atc gag atc ggt gaa acc gag gga acc ttc atc gtg gac agc 4722
 Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser
 1155 1160 1165
 gtg gag ctt ctc ttg atg gag gaa taa tga gatctatcga ttctagaagg 4772
 Val Glu Leu Leu Leu Met Glu Glu
 1170 1175
 cctgaattct gcatgcgttt ggacgtatgc tcattcaggt tggagccaat ttggttgatg 4832
 tgtgtgcgag ttcttgcgag tctgatgaga catctctgta ttgtgtttct ttccccagtg 4892
 ttttctgtac ttgtgtaatc ggctaatacgc caacagattc ggcgatgaat aaatgagaaa 4952
 taaattgttc tgattttgag tgcaaaaaaa aaggaatt 4990

<210> 8
 <211> 1177
 <212> PRT
 <213> artificial

<220>
 <223> Synthetic Construct

<400> 8

Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
1 5 10 15

Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly
20 25 30

Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
35 40 45

Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
50 55 60

Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
65 70 75 80

Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
85 90 95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu

```

100          105          110
Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu
115          120          125
Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala
130          135          140
Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val
145          150          155          160
Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser
165          170          175
Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg
180          185          190
Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp His Ala Val
195          200          205
Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg
210          215          220
Asp Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val
225          230          235          240
Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr Tyr Pro
245          250          255
Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val
260          265          270
Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu
275          280          285
Gly Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr
290          295          300
Ile Tyr Thr Asp Ala His Arg Gly Glu Tyr Tyr Trp Ser Gly His Gln
305          310          315          320
Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro
325          330          335
Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
340          345          350
Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
355          360          365
Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp
370          375          380
Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val
385          390          395          400
Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln
405          410          415
Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His

```

420 425 430
 Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
 435 440 445
 Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn
 450 455 460
 Ile Ile Ala Ser Asp Ser Ile Thr Gln Ile Pro Leu Val Lys Ala His
 465 470 475 480
 Thr Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro Gly Phe Thr Gly
 485 490
 Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr Ile
 500 505 510
 Val Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg Ala Arg Ile Arg
 515 520 525
 Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu
 530 535 540
 Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly Asp Pro
 545 550 555 560
 Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr
 565 570 575
 Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe Ser
 580 585 590
 Ser Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu Ile Pro Val Thr
 595 600 605
 Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala Val
 610 615 620
 Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val
 625 630 635 640
 Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser
 645 650 655
 Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
 660 665 670
 His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn
 675 680 685
 Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr
 690 695 700
 Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
 705 710 715 720
 Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
 725 730 735
 Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg

```

740              745              750

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Ser Ile Arg Tyr
755              760              765

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
770              775              780

Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg
785              790              795              800

Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg
805              810              815

Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile
820              825              830

Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile
835              840              845

Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu
850              855              860

Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys
865              870              875              880

Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu
885              890              895

Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe
900              905              910

Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met
915              920              925

Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu
930              935              940

Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu
945              950              955              960

Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn
965              970              975

Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val
980              985              990

Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu
995              1000              1005

Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val
1010              1015              1020

Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu
1025              1030              1035

Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn
1040              1045              1050

Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr

```

1055	1060	1065
Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu 1070 1075 1080		
Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu 1085 1090 1095		
Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys 1100 1105 1110		
Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg 1115 1120 1125		
Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys 1130 1135 1140		
Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile 1145 1150 1155		
Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu 1160 1165 1170		
Leu Met Glu Glu 1175		