COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF TUMOR

RELATED APPLICATIONS

This application is a continuation-in-part of, and claims priority under 35 U.S.C. §120 to, U.S. Application Serial No. 10/177,488 filed 6/19/2002, which claims priority under 35 U.S.C. §119(e) to, U.S. Provisional Application Serial Nos. 60/299,500 filed 6/20/2001, 60/301,880 filed 6/29/2001, and 60/323,268 filed 9/18/2001, the disclosures of which are herein incorporated by reference.

FIELD OF THE INVENTION

The present invention is directed to compositions of matter useful for the diagnosis and treatment of tumor in mammals and to methods of using those compositions of matter for the same.

BACKGROUND OF THE INVENTION

Malignant tumors (cancers) are the second leading cause of death in the United States, after heart disease (Boring et al., *CA Cancel J. Clin.* 43:7 (1993)). Cancer is characterized by the increase in the number of abnormal, or neoplastic, cells derived from a normal tissue which proliferate to form a tumor mass, the invasion of adjacent tissues by these neoplastic tumor cells, and the generation of malignant cells which eventually spread via the blood or lymphatic system to regional lymph nodes and to distant sites via a process called metastasis. In a cancerous state, a cell proliferates under conditions in which normal cells would not grow. Cancer manifests itself in a wide variety of forms, characterized by different degrees of invasiveness and aggressiveness.

In attempts to discover effective cellular targets for cancer diagnosis and therapy, researchers have sought to identify transmembrane or otherwise membrane-associated polypeptides that are specifically expressed on the surface of one or more particular type(s) of cancer cell as compared to on one or more normal non-cancerous cell(s). Often, such membrane-associated polypeptides are more abundantly expressed on the surface of the cancer cells as compared to on the surface of the non-cancerous cells. The identification of such tumor-associated cell surface antigen polypeptides has given rise to the ability to specifically target cancer cells for destruction via antibody-based therapies. In this regard, it is noted that antibody-based therapy has proved very effective in the treatment of certain cancers. For example, HERCEPTIN® and RITUXAN® (both from Genentech Inc., South San Francisco, California) are antibodies that have been used successfully to treat breast cancer and non-Hodgkin's lymphoma, respectively. More specifically, HERCEPTIN® is a recombinant DNA-derived humanized monoclonal antibody that selectively binds to the extracellular domain of the human epidermal growth factor receptor 2 (HER2) proto-oncogene. HER2 protein overexpression is observed in 25-30% of primary breast cancers. RITUXAN® is a genetically engineered chimeric murine/human monoclonal antibody directed against the CD20 antigen found on the surface of normal and malignant B lymphocytes. Both these antibodies are recombinantly produced in CHO cells.

In other attempts to discover effective cellular targets for cancer diagnosis and therapy, researchers have sought to identify (1) non-membrane-associated polypeptides that are specifically produced by one or more particular type(s) of cancer cell(s) as compared to by one or more particular type(s) of non-cancerous normal cell(s), (2) polypeptides that are produced by cancer cells at an expression level that is significantly higher than that of one or more normal non-cancerous cell(s), or (3) polypeptides whose expression is specifically limited to only a single (or very limited number of different) tissue type(s) in both the cancerous and

(e.g., normal prostate and prostate tumor tissue). Such polypeptides may remain intracellularly located or may be secreted by the cancer cell. Moreover, such polypeptides may be expressed not by the cancer cell itself, but rather by cells which produce and/or secrete polypeptides having a potentiating or growth-enhancing effect on cancer cells. Such secreted polypeptides are often proteins that provide cancer cells with a growth advantage over normal cells and include such things as, for example, angiogenic factors, cellular adhesion factors, growth factors, and the like. Identification of antagonists of such non-membrane associated polypeptides would be expected to serve as effective therapeutic agents for the treatment of such cancers. Furthermore, identification of the expression pattern of such polypeptides would be useful for the diagnosis and treatment of particular cancers in mammals.

The ability to modulate gene expression in a mammal is still difficult. Traditionally, it has been done using tools such viral vectors that will express a polypeptide that the host is lacking. The abiltity to introduce a viral vector into a host that will reduce gene expression has not been perfected. Repression of gene expression has traditionally been done in mammals in a "knockout" fashion, where the test mammal, usually a mouse, has the gene ablated or "knocked out" through the technique of homolgous recombination. This techinique in mice is slow, laborious and difficult, and impractical in humans. Therefore, Applicants turned to a vector system which will express an interfering RNA (si RNA) to reduce gene expression.

siRNAs have proven useful as a tool in studies of modulating gene expression where traditional antagonists such as small molecules or antibodies have failed. (Shi Y., *Trends in Genetics* 19(1):9-12 (2003)). In vitro synthesized, double stranded RNAs that are 21 to 23 nucleotides in length can act as interfering RNAs (iRNAs) and can specifically inhibit gene expression (Fire A., *Trends in Genetics* 391; 806-810 (1999)). These iRNAs act by mediating degradation of their target RNAs. Since they are under 30 nuclotides in length, however they do not trigger a cell antiviral defense mechanism. Such mechanisms include interferon production, and a general shutdown of host cell protein synthesis. Practically, siRNAs can by synthesized and then cloned into DNA vectors. Such vectors can be transfected and made to express the siRNA at high levels and/or in a tissue specific manner. The high level of siRNA expression is used to "knockdown" or significantly reduce the amount of protein produced in a cell, and thus it is useful in experiments where overexpression of a protein is believed to be linked to a disorders such as cancer. Despite advances in mammalian cancer therapy, there is a great need for therapeutic agents capable of effectively inhibiting neoplastic cell growth through reduction of gene expression. Accordingly, it is an objective of the present invention to identify a system that will modulate gene expression.

Improving the delivery of drugs and other agentsto target cells, tissues and tumors to achieve maximal efficacy and minimal toxicity has been the focus of considerable research for many years. Though many attempts have been made to develop effective methods for importing biologically active molecules into cells, both *in vivo* and *in vitro*, none has proved to be entirely satisfactory. Optimizing the association of the drug with its intracellular target, while minimizing intercellular redistribution of the drug, e.g. to neighboring cells, is often difficult or inefficient.

Most agents currently administered to a patient parenterally are not targeted, resulting in systemic delivery of the agent to cells and tissues of the body where it is unnecessary, and often undesirable. This may

result in adverse drug side effects, and often limits the dose of a drug (e.g., chemotherapeutic (anti-cancer), cytotoxic, enzyme inhibitor agents and antiviral or antimicrobial drugs) that can be administered. By comparison, although oral administration of drugs is considered to be a convenient and economical mode of administration, it shares the same concerns of non-specific toxicity to unaffected cells once the drug has been absorbed into the systemic circulation. Further complications involve problems with oral bioavailability and residence of drug in the gut leading to additional exposure of gut to the drug and hence risk of gut toxicities. Accordingly, a major goal has been to develop methods for specifically targeting agents to cells and tissues. The benefits of such treatment include avoiding the general physiological effects of inappropriate delivery of such agents to other cells and tissues, such as uninfected cells. Intracellular targeting may be achieved by methods, compounds and formulations which allow accumulation or retention of biologically active agents, i.e. active metabolites, inside cells.

Monoclonal antibody therapy has been established for the targeted treatment of patients with cancer, immunological and angiogenic disorders.

There are a variety of useful toxins available. For example, the auristatin peptides, auristain E (AE) and monomethylauristatin (MMAE), synthetic analogs of dolastatin, were conjugated to: (i) chimeric monoclonal antibodies cBR96 (specific to Lewis Y on carcinomas); (ii) cAC10 which is specific to CD30 on hematological malignancies (Klussman, et al (2004), Bioconjugate Chemistry 15(4):765-773; Doronina et al (2003) Nature Biotechnology 21(7):778-784; "Monomethylvaline Compounds Capable of Conjugation to Ligands"; Francisco et al (2003) Blood 102(4):1458-1465; US 2004/0018194; (iii) anti-CD20 antibodies such as rituxan (WO 04/032828) for the treatment of CD20-expressing cancers and immune disorders; (iv) anti-EphB2 antibodies 2H9 and anti-IL-8 for treatment of colorectal cancer (Mao, et al (2004) Cancer Research 64(3):781-788); (v) E-selectin antibody (Bhaskar et al (2003) Cancer Res. 63:6387-6394); and (vi) other anti-CD30 antibodies (WO 03/043583). Monomethylauristatin (MMAE) has also been conjugated to 2H9, an antibody against EphB2R which is a type 1 TM tyrosine kinase receptor with close homology between mouse and human, and is over-expressed in colorectal cancer cells (Mao et al (2004) Cancer Res. 64:781-788).

Monomethylauristatin MMAF, a variant of auristatin E (MMAE) with a phenylalanine at the C-terminus (US 5767237; US 6124431), has been reported to be less potent than MMAE, but more potent when conjugated to monoclonal antibodies (Senter et al, Proceedings fo the American Association for Cancer Research, Volume 45, Abstract Number 623, presented March 28, 2004). Auristatin F phenylene diamine (AFP); a phenylalanine variant of MMAE was linked to an anti-CD70 mAb, 1F6, through the C-terminus of 1F6 via a phenylene diamine spacer (Law et al, Proceedings of the American Association for Cancer Research, Volume 45, Abstract Number 625, presented March 28, 2004)

Despite the above identified advances in mammalian cancer therapy, there is a great need for additional diagnostic and therapeutic agents capable of detecting the presence of tumor in a mammal and for effectively inhibiting neoplastic cell growth, respectively. Accordingly, it is an objective of the present invention to identify: (1) cell membrane-associated polypeptides that are more abundantly expressed on one or more type(s) of cancer cell(s) as compared to on normal cells or on other different cancer cells, (2) non-membrane-associated

polypeptides that are specifically produced by one or more particular type(s) of cancer cell(s) (or by other cells that produce polypeptides having a potentiating effect on the growth of cancer cells) as compared to by one or more particular type(s) of non-cancerous normal cell(s), (3) non-membrane-associated polypeptides that are produced by cancer cells at an expression level that is significantly higher than that of one or more normal non-cancerous cell(s), or (4) polypeptides whose expression is specifically limited to only a single (or very limited number of different) tissue type(s) in both a cancerous and non-cancerous state (e.g., normal prostate and prostate tumor tissue), and to use those polypeptides, and their encoding nucleic acids, to produce compositions of matter useful in the therapeutic treatment and diagnostic detection of cancer in mammals. It is also an objective of the present invention to identify cell membrane-associated, secreted or intracellular polypeptides whose expression is limited to a single or very limited number of tissues, and to use those polypeptides, and their encoding nucleic acids, to produce compositions of matter useful in the therapeutic treatment and diagnostic detection of cancer in mammals.

SUMMARY OF THE INVENTION

A. Embodiments

In the present specification, Applicants describe for the first time the identification of various cellular polypeptides (and their encoding nucleic acids or fragments thereof) which are expressed to a greater degree on the surface of or by one or more types of cancer cell(s) as compared to on the surface of or by one or more types of normal non-cancer cells. Alternatively, such polypeptides are expressed by cells which produce and/or secrete polypeptides having a potentiating or growth-enhancing effect on cancer cells. Again alternatively, such polypeptides may not be overexpressed by tumor cells as compared to normal cells of the same tissue type, but rather may be specifically expressed by both tumor cells and normal cells of only a single or very limited number of tissue types (preferably tissues which are not essential for life, e.g., prostate, etc.). All of the above polypeptides are herein referred to as **T**umor-associated **A**ntigenic **T**arget polypeptides ("TAT" polypeptides) and are expected to serve as effective targets for cancer therapy and diagnosis in mammals.

Accordingly, in one embodiment of the present invention, the invention provides an isolated nucleic acid molecule having a nucleotide sequence that encodes a tumor-associated antigenic target polypeptide or fragment thereof (a "TAT" polypeptide).

In certain aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% nucleic acid sequence identity, to (a) a DNA molecule encoding a full-length TAT polypeptide having an amino acid sequence as disclosed herein, a TAT polypeptide amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane TAT polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length TAT polypeptide amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least

about 80% nucleic acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% nucleic acid sequence identity, to (a) a DNA molecule comprising the coding sequence of a full-length TAT polypeptide cDNA as disclosed herein, the coding sequence of a TAT polypeptide lacking the signal peptide as disclosed herein, the coding sequence of an extracellular domain of a transmembrane TAT polypeptide, with or without the signal peptide, as disclosed herein or the coding sequence of any other specifically defined fragment of the full-length TAT polypeptide amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In further aspects, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% nucleic acid sequence identity, to (a) a DNA molecule that encodes the same mature polypeptide encoded by the full-length coding region of any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

Another aspect of the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a TAT polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptide(s) are disclosed herein. Therefore, soluble extracellular domains of the herein described TAT polypeptides are contemplated.

In other aspects, the present invention is directed to isolated nucleic acid molecules which hybridize to (a) a nucleotide sequence encoding a TAT polypeptide having a full-length amino acid sequence as disclosed herein, a TAT polypeptide amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane TAT polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length TAT polypeptide amino acid sequence as disclosed herein, or (b) the complement of the nucleotide sequence of (a). In this regard, an embodiment of the present invention is directed to fragments of a full-length TAT polypeptide coding sequence, or the complement thereof, as disclosed herein, that may find use as, for example, hybridization probes useful as, for example, diagnostic probes, antisense oligonucleotide probes, or for encoding fragments of a full-length TAT polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-TAT polypeptide antibody, a TAT binding oligopeptide or other small organic molecule that binds to a TAT polypeptide. Such nucleic acid fragments are usually at least about 5 nucleotides in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, or 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted

that novel fragments of a TAT polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the TAT polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which TAT polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such novel fragments of TAT polypeptide-encoding nucleotide sequences are contemplated herein. Also contemplated are the TAT polypeptide fragments encoded by these nucleotide molecule fragments, preferably those TAT polypeptide fragments that comprise a binding site for an anti-TAT antibody, a TAT binding oligopeptide or other small organic molecule that binds to a TAT polypeptide.

In another embodiment, the invention provides isolated TAT polypeptides encoded by any of the isolated nucleic acid sequences hereinabove identified.

In a certain aspect, the invention concerns an isolated TAT polypeptide, comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% amino acid sequence identity, to a TAT polypeptide having a full-length amino acid sequence as disclosed herein, a TAT polypeptide amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane TAT polypeptide protein, with or without the signal peptide, as disclosed herein, an amino acid sequence encoded by any of the nucleic acid sequences disclosed herein or any other specifically defined fragment of a full-length TAT polypeptide amino acid sequence as disclosed herein.

In a further aspect, the invention concerns an isolated TAT polypeptide comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% amino acid sequence identity, to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

In a specific aspect, the invention provides an isolated TAT polypeptide without the N-terminal signal sequence and/or without the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the TAT polypeptide and recovering the TAT polypeptide from the cell culture.

Another aspect of the invention provides an isolated TAT polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the TAT polypeptide and recovering the TAT polypeptide from the cell culture.

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cells comprising any such vector are also provided. By way of example, the host cells may be CHO cells*E. coli* cells, or yeast cells. A process for producing any of the

herein described polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

In other embodiments, the invention provides isolated chimeric polypeptides comprising any of the herein described TAT polypeptides fused to a heterologous (non-TAT) polypeptide. Example of such chimeric molecules comprise any of the herein described TAT polypeptides fused to a heterologous polypeptide such as, for example, an epitope tag sequence or a Fc region of an immunoglobulin.

In another embodiment, the invention provides an antibody which binds, preferably specifically, to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, antibody fragment, chimeric antibody, humanized antibody, single-chain antibody or antibody that competitively inhibits the binding of an anti-TAT polypeptide antibody to its respective antigenic epitope. Antibodies of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies of the present invention may optionally be produced in CHO cells or bacterial cells and preferably induce death of a cell to which they bind. For diagnostic purposes, the antibodies of the present invention may be detectably labeled, attached to a solid support, or the like.

In another embodiment of the present invention comprising antibody conjugated to a cytotoxin, the toxin may be selected from any of the toxins known in the art. In one embodiment, the toxin is linked to the antibody by a peptide or petidomimetic linker. In one embodiment, the toxin is linked to the antibody via a linker comprising valine-citrulline (-vc-). In one embodiment the toxin is MMAE. In one embodiment the toxin is MMAE. In one embodiment, the toxin is an auristatin peptide.

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described antibodies. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells*E. coli* cells, or yeast cells. A process for producing any of the herein described antibodies is further provided and comprises culturing host cells under conditions suitable for expression of the desired antibody and recovering the desired antibody from the cell culture.

In another embodiment, the invention provides oligopeptides ("TAT binding oligopeptides") which bind, preferably specifically, to any of the above or below described TAT polypeptides. Optionally, the TAT binding oligopeptides of the present invention may be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The TAT binding oligopeptides of the present invention may optionally be produced in CHO cells or bacterial cells and preferably induce death of a cell to which they bind. For diagnostic purposes, the TAT binding oligopeptides of the present invention may be detectably labeled, attached to a solid support, or the like.

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described TAT binding oligopeptides. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells*E. coli* cells, or yeast cells. A process for

producing any of the herein described TAT binding oligopeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired oligopeptide and recovering the desired oligopeptide from the cell culture.

In another embodiment, the invention provides small organic molecules ("TAT binding organic molecules") which bind, preferably specifically, to any of the above or below described TAT polypeptides. Optionally, the TAT binding organic molecules of the present invention may be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The TAT binding organic molecules of the present invention preferably induce death of a cell to which they bind. For diagnostic purposes, the TAT binding organic molecules of the present invention may be detectably labeled, attached to a solid support, or the like.

In a still further embodiment, the invention concerns a composition of matter comprising a TAT polypeptide as described herein, a chimeric TAT polypeptide as described herein, an anti-TAT antibody as described herein, a TAT binding oligopeptide as described herein, or a TAT binding organic molecule as described herein, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

In a still further embodiment, the invention concerns a composition of matter comprising a TAT polypeptide as described herein, a chimeric TAT polypeptide as described herein, an anti-TAT antibody as described herein, a TAT binding oligopeptide as described herein, a TAT binding interfereing RNA (siRNA) or a TAT binding organic molecule as described herein, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier. Specifically the TAT polypeptide is TAT188, also referred to herein as E16 or as TAT188(E16). The anti-TAT188 (also referred to as anti-E16 or anti-TAT188(E16)) antibodies of one embodiment of the invention include without limitation 3B5, 12B9, and 12G12, as described herein. The siRNA of the embodiment includes siTAT188 (also referred to as siE16 or siTAT188(E16 or TAT188 siRNA or E16 siRNA or TAT188(E16) siRNA)) as described herein.

In yet another embodiment, the invention concerns an article of manufacture comprising a container and a composition of matter contained within the container, wherein the composition of matter may comprise a TAT polypeptide as described herein, a chimeric TAT polypeptide as described herein, an anti-TAT antibody as described herein, a TAT binding oligopeptide as described herein, or a TAT binding organic molecule as described herein. The article may further optionally comprise a label affixed to the container, or a package insert included with the container, that refers to the use of the composition of matter for the therapeutic treatment or diagnostic detection of a tumor.

Another embodiment of the present invention is directed to the use of a TAT polypeptide as described herein, a chimeric TAT polypeptide as described herein, an anti-TAT polypeptide antibody as described herein, a TAT binding oligopeptide as described herein, or a TAT binding organic molecule as described herein, for the preparation of a medicament useful in the treatment of a condition which is responsive to the TAT polypeptide, chimeric TAT polypeptide, anti-TAT polypeptide antibody, TAT binding oligopeptide, or TAT binding organic molecule.

B. Additional Embodiments

Another embodiment of the present invention is directed to a method for inhibiting the growth of a cell that expresses a TAT polypeptide, wherein the method comprises contacting the cell with an antibody, an oligopeptide or a small organic molecule that binds to the TAT polypeptide, and wherein the binding of the antibody, oligopeptide or organic molecule to the TAT polypeptide causes inhibition of the growth of the cell expressing the TAT polypeptide. In preferred embodiments, the cell is a cancer cell and binding of the antibody, oligopeptide or organic molecule to the TAT polypeptide causes death of the cell expressing the TAT polypeptide. Optionally, the antibody is a monoclonal antibody, antibody fragment, chimeric antibody, humanized antibody, or single-chain antibody. Antibodies, TAT binding oligopeptides and TAT binding organic molecules employed in the methods of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies and TAT binding oligopeptides employed in the methods of the present invention may optionally be produced in CHO cells or bacterial cells.

Yet another embodiment of the present invention is directed to a method of therapeutically treating a mammal having a cancerous tumor comprising cells that express a TAT polypeptide, wherein the method comprises administering to the mammal a therapeutically effective amount of an antibody, an oligopeptide or a small organic molecule that binds to the TAT polypeptide, thereby resulting in the effective therapeutic treatment of the tumor. Optionally, the antibody is a monoclonal antibody, antibody fragment, chimeric antibody, humanized antibody, or single-chain antibody. Antibodies, TAT binding oligopeptides and TAT binding organic molecules employed in the methods of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies and oligopeptides employed in the methods of the present invention may optionally be produced in CHO cells or bacterial cells.

Yet another embodiment of the present invention is directed to a method of determining the presence of a TAT polypeptide in a sample suspected of containing the TAT polypeptide, wherein the method comprises exposing the sample to an antibody, oligopeptide or small organic molecule that binds to the TAT polypeptide and determining binding of the antibody, oligopeptide or organic molecule to the TAT polypeptide in the sample, wherein the presence of such binding is indicative of the presence of the TAT polypeptide in the sample. Optionally, the sample may contain cells (which may be cancer cells) suspected of expressing the TAT polypeptide. The antibody, TAT binding oligopeptide or TAT binding organic molecule employed in the method may optionally be detectably labeled, attached to a solid support, or the like.

A further embodiment of the present invention is directed to a method of diagnosing the presence of a tumor in a mammal, wherein the method comprises detecting the level of expression of a gene encoding a TAT polypeptide (a) in a test sample of tissue cells obtained from said mammal, and (b) in a control sample of known normal non-cancerous cells of the same tissue origin or type, wherein a higher level of expression of

the TAT polypeptide in the test sample, as compared to the control sample, is indicative of the presence of tumor in the mammal from which the test sample was obtained.

Another embodiment of the present invention is directed to a method of diagnosing the presence of a tumor in a mammal, wherein the method comprises (a) contacting a test sample comprising tissue cells obtained from the mammal with an antibody, oligopeptide or small organic molecule that binds to a TAT polypeptide and (b) detecting the formation of a complex between the antibody, oligopeptide or small organic molecule and the TAT polypeptide in the test sample, wherein the formation of a complex is indicative of the presence of a tumor in the mammal. Optionally, the antibody, TAT binding oligopeptide or TAT binding organic molecule employed is detectably labeled, attached to a solid support, or the like, and/or the test sample of tissue cells is obtained from an individual suspected of having a cancerous tumor.

Yet another embodiment of the present invention is directed to a method for treating or preventing a cell proliferative disorder associated with altered, preferably increased, expression or activity of a TAT polypeptide, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a TAT polypeptide. Preferably, the cell proliferative disorder is cancer and the antagonist of the TAT polypeptide is an anti-TAT polypeptide antibody, TAT binding oligopeptide, TAT binding organic molecule or antisense oligonucleotide. Effective treatment or prevention of the cell proliferative disorder may be a result of direct killing or growth inhibition of cells that express a TAT polypeptide or by antagonizing the cell growth potentiating activity of a TAT polypeptide.

Yet another embodiment of the present invention is directed to a method of binding an antibody, oligopeptide or small organic molecule to a cell that expresses a TAT polypeptide, wherein the method comprises contacting a cell that expresses a TAT polypeptide with said antibody, oligopeptide or small organic molecule under conditions which are suitable for binding of the antibody, oligopeptide or small organic molecule to said TAT polypeptide and allowing binding therebetween.

Other embodiments of the present invention are directed to the use of (a) a TAT polypeptide, (b) a nucleic acid encoding a TAT polypeptide or a vector or host cell comprising that nucleic acid, (c) an anti-TAT polypeptide antibody, (d) a TAT-binding oligopeptide, or (e) a TAT-binding small organic molecule in the preparation of a medicament useful for (i) the therapeutic treatment or diagnostic detection of a cancer or tumor, or (ii) the therapeutic treatment or prevention of a cell proliferative disorder.

Another embodiment of the present invention is directed to a method for inhibiting the growth of a cancer cell, wherein the growth of said cancer cell is at least in part dependent upon the growth potentiating effect(s) of a TAT polypeptide (wherein the TAT polypeptide may be expressed either by the cancer cell itself or a cell that produces polypeptide(s) that have a growth potentiating effect on cancer cells), wherein the method comprises contacting the TAT polypeptide with an antibody, an oligopeptide or a small organic molecule that binds to the TAT polypeptide, thereby antagonizing the growth-potentiating activity of the TAT polypeptide and, in turn, inhibiting the growth of the cancer cell. Preferably the growth of the cancer cell is completely inhibited. Even more preferably, binding of the antibody, oligopeptide or small organic molecule to the TAT polypeptide induces the death of the cancer cell. Optionally, the antibody is a monoclonal

antibody, antibody fragment, chimeric antibody, humanized antibody, or single-chain antibody. Antibodies, TAT binding oligopeptides and TAT binding organic molecules employed in the methods of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies and TAT binding oligopeptides employed in the methods of the present invention may optionally be produced in CHO cells or bacterial cells.

Yet another embodiment of the present invention is directed to a method of therapeutically treating a tumor in a mammal, wherein the growth of said tumor is at least in part dependent upon the growth potentiating effect(s) of a TAT polypeptide, wherein the method comprises administering to the mammal a therapeutically effective amount of an antibody, an oligopeptide or a small organic molecule that binds to the TAT polypeptide, thereby antagonizing the growth potentiating activity of said TAT polypeptide and resulting in the effective therapeutic treatment of the tumor. Optionally, the antibody is a monoclonal antibody, antibody fragment, chimeric antibody, humanized antibody, or single-chain antibody. Antibodies, TAT binding oligopeptides and TAT binding organic molecules employed in the methods of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies and oligopeptides employed in the methods of the present invention may optionally be produced in CHO cells or bacterial cells.

C. Further Additional Embodiments

In yet further embodiments, the invention is directed to the following set of potential claims for this application:

1. Isolated nucleic acid having a nucleotide sequence that has at least 80% nucleic acid sequence identity to:

(a) a DNA molecule encoding the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) a DNA molecule encoding the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) a DNA molecule encoding an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) a DNA molecule encoding an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78);

(f) the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(g) the complement of (a), (b), (c), (d), (e) or (f).

2. Isolated nucleic acid having:

(a) a nucleotide sequence that encodes the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) a nucleotide sequence that encodes the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) a nucleotide sequence that encodes an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) a nucleotide sequence that encodes an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78);

(f) the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(g) the complement of (a), (b), (c), (d), (e) or (f).

3. Isolated nucleic acid that hybridizes to:

(a) a nucleic acid that encodes the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) a nucleic acid that encodes the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) a nucleic acid that encodes an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) a nucleic acid that encodes an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78);

(f) the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(g) the complement of (a), (b), (c), (d), (e) or (f).

4. The nucleic acid of Claim 3, wherein the hybridization occurs under stringent conditions.

5. The nucleic acid of Claim 3 which is at least about 5 nucleotides in length.

6. An expression vector comprising the nucleic acid of Claim 1, 2 or 3.

7. The expression vector of Claim 6, wherein said nucleic acid is operably linked to control sequences recognized by a host cell transformed with the vector.

8. A host cell comprising the expression vector of Claim 7.

9. The host cell of Claim 8 which is a CHO cell, an *E. coli* cell or a yeast cell.

10. A process for producing a polypeptide comprising culturing the host cell of Claim 8 under conditions suitable for expression of said polypeptide and recovering said polypeptide from the cell culture.

11. An isolated polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

12. An isolated polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

13. A chimeric polypeptide comprising the polypeptide of Claim 11 or 12 fused to a heterologous polypeptide.

14. The chimeric polypeptide of Claim 13, wherein said heterologous polypeptide is an epitope tag sequence or an Fc region of an immunoglobulin.

15. An isolated antibody that binds to a polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

16. An isolated antibody that binds to a polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

17. The antibody of Claim 15 or 16 which is a monoclonal antibody.

18. The antibody of Claim 15 or 16 which is an antibody fragment.

19. The antibody of Claim 15 or 16 which is a chimeric or a humanized antibody.

20. The antibody of Claim 15 or 16 which is conjugated to a growth inhibitory agent.

21. The antibody of Claim 15 or 16 which is conjugated to a cytotoxic agent.

22. The antibody of Claim 21, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

23. The antibody of Claim 21, wherein the cytotoxic agent is a toxin.

24. The antibody of Claim 23, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

25. The antibody of Claim 23, wherein the toxin is a maytansinoid.

26. The antibody of Claim 15 or 16 which is produced in bacteria.

27. The antibody of Claim 15 or 16 which is produced in CHO cells.

28. The antibody of Claim 15 or 16 which induces death of a cell to which it binds.

29. The antibody of Claim 15 or 16 which is detectably labeled.

30. An isolated nucleic acid having a nucleotide sequence that encodes the antibody of Claim 15 or 16.

31. An expression vector comprising the nucleic acid of Claim 30 operably linked to control sequences recognized by a host cell transformed with the vector.

32. A host cell comprising the expression vector of Claim 31.

33. The host cell of Claim 32 which is a CHO cell, an *E. coli* cell or a yeast cell.

34. A process for producing an antibody comprising culturing the host cell of Claim 32 under conditions suitable for expression of said antibody and recovering said antibody from the cell culture.

35. An isolated oligopeptide that binds to a polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

36. An isolated oligopeptide that binds to a polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

37. The oligopeptide of Claim 35 or 36 which is conjugated to a growth inhibitory agent.

38. The oligopeptide of Claim 35 or 36 which is conjugated to a cytotoxic agent.

39. The oligopeptide of Claim 38, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

40. The oligopeptide of Claim 38, wherein the cytotoxic agent is a toxin.

41. The oligopeptide of Claim 40, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

42. The oligopeptide of Claim 40, wherein the toxin is a maytansinoid.

43. The oligopeptide of Claim 35 or 36 which induces death of a cell to which it binds.

44. The oligopeptide of Claim 35 or 36 which is detectably labeled.

45. A TAT binding organic molecule that binds to a polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

46. The organic molecule of Claim 45 that binds to a polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

47. The organic molecule of Claim 45 or 46 which is conjugated to a growth inhibitory agent.

48. The organic molecule of Claim 45 or 46 which is conjugated to a cytotoxic agent.

49. The organic molecule of Claim 48, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

50. The organic molecule of Claim 48, wherein the cytotoxic agent is a toxin.

51. The organic molecule of Claim 50, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

52. The organic molecule of Claim 50, wherein the toxin is a maytansinoid.

53. The organic molecule of Claim 45 or 46 which induces death of a cell to which it binds.

54. The organic molecule of Claim 45 or 46 which is detectably labeled.

55. A composition of matter comprising:

(a) the polypeptide of Claim 11;

(b) the polypeptide of Claim 12;

(c) the chimeric polypeptide of Claim 13;

(d) the antibody of Claim 15;

(e) the antibody of Claim 16;

(f) the oligopeptide of Claim 35;

(g) the oligopeptide of Claim 36;

(h) the TAT binding organic molecule of Claim 45; or

(i) the TAT binding organic molecule of Claim 46; in combination with a carrier.

56. The composition of matter of Claim 55, wherein said carrier is a pharmaceutically acceptable carrier.

57. An article of manufacture comprising:

(a) a container; and

(b) the composition of matter of Claim 55 contained within said container.

58. The article of manufacture of Claim 57 further comprising a label affixed to said container, or a package insert included with said container, referring to the use of said composition of matter for the therapeutic treatment of or the diagnostic detection of a cancer.

59. A method of inhibiting the growth of a cell that expresses a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising contacting said cell with an antibody, oligopeptide or organic molecule that binds to said protein, the binding of said antibody, oligopeptide or organic molecule to said protein thereby causing an inhibition of growth of said cell.

60. The method of Claim 59, wherein said antibody is a monoclonal antibody.

61. The method of Claim 59, wherein said antibody is an antibody fragment.

62. The method of Claim 59, wherein said antibody is a chimeric or a humanized antibody.

63. The method of Claim 59, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.

64. The method of Claim 59, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.

65. The method of Claim 64, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

66. The method of Claim 64, wherein the cytotoxic agent is a toxin.

67. The method of Claim 66, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

68. The method of Claim 66, wherein the toxin is a maytansinoid.

69. The method of Claim 59, wherein said antibody is produced in bacteria.

70. The method of Claim 59, wherein said antibody is produced in CHO cells.

71. The method of Claim 59, wherein said cell is a cancer cell.

72. The method of Claim 71, wherein said cancer cell is further exposed to radiation treatment or a chemotherapeutic agent.

73. The method of Claim 71, wherein said cancer cell is selected from the group consisting of a breast cancer cell, a colorectal cancer cell, a lung cancer cell, an ovarian cancer cell, a central nervous system cancer cell, a liver cancer cell, a bladder cancer cell, a pancreatic cancer cell, a cervical cancer cell, a melanoma cell and a leukemia cell.

74. The method of Claim 71, wherein said protein is more abundantly expressed by said cancer cell as compared to a normal cell of the same tissue origin.

75. The method of Claim 59 which causes the death of said cell.

76. The method of Claim 59, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

77. A method of therapeutically treating a mammal having a cancerous tumor comprising cells that express a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising administering to said mammal a therapeutically effective amount of an antibody, oligopeptide or organic molecule that binds to said protein, thereby effectively treating said mammal.

78. The method of Claim 77, wherein said antibody is a monoclonal antibody.

79. The method of Claim 77, wherein said antibody is an antibody fragment.

80. The method of Claim 77, wherein said antibody is a chimeric or a humanized antibody.

81. The method of Claim 77, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.

82. The method of Claim 77, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.

83. The method of Claim 82, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

84. The method of Claim 82, wherein the cytotoxic agent is a toxin.

85. The method of Claim 84, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

86. The method of Claim 84, wherein the toxin is a maytansinoid.

87. The method of Claim 77, wherein said antibody is produced in bacteria.

88. The method of Claim 77, wherein said antibody is produced in CHO cells.

89. The method of Claim 77, wherein said tumor is further exposed to radiation treatment or a chemotherapeutic agent.

90. The method of Claim 77, wherein said tumor is a breast tumor, a colorectal tumor, a lung tumor, an ovarian tumor, a central nervous system tumor, a liver tumor, a bladder tumor, a pancreatic tumor, or a cervical tumor.

91. The method of Claim 77, wherein said protein is more abundantly expressed by the cancerous cells of said tumor as compared to a normal cell of the same tissue origin.

92. The method of Claim 77, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of

Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

93. A method of determining the presence of a protein in a sample suspected of containing said protein, wherein said protein has at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising exposing said sample to an antibody, oligopeptide or organic molecule that binds to said protein and determining binding of said antibody, oligopeptide or organic molecule to said protein in said sample, wherein binding of the antibody, oligopeptide or organic molecule to said protein is indicative of the presence of said protein in said sample.

94. The method of Claim 93, wherein said sample comprises a cell suspected of expressing said protein.

95. The method of Claim 94, wherein said cell is a cancer cell.

96. The method of Claim 93, wherein said antibody, oligopeptide or organic molecule is detectably labeled.

97. The method of Claim 93, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence

shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

98. A method of diagnosing the presence of a tumor in a mammal, said method comprising determining the level of expression of a gene encoding a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), in a test sample of tissue cells obtained from said mammal and in a control sample of known normal cells of the same tissue origin, wherein a higher level of expression of said protein in the test sample, as compared to the control sample, is indicative of the presence of tumor in the mammal from which the test sample was obtained.

99. The method of Claim 98, wherein the step of determining the level of expression of a gene encoding said protein comprises employing an oligonucleotide in an *in situ* hybridization or RT-PCR analysis.

100. The method of Claim 98, wherein the step determining the level of expression of a gene encoding said protein comprises employing an antibody in an immunohistochemistry or Western blot analysis.

101. The method of Claim 98, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

102. A method of diagnosing the presence of a tumor in a mammal, said method comprising contacting a test sample of tissue cells obtained from said mammal with an antibody, oligopeptide or organic molecule that binds to a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), and detecting the formation of a complex between said antibody, oligopeptide or organic molecule and said protein in the test sample, wherein the formation of a complex is indicative of the presence of a tumor in said mammal.

103. The method of Claim 102, wherein said antibody, oligopeptide or organic molecule is detectably labeled.

104. The method of Claim 102, wherein said test sample of tissue cells is obtained from an individual suspected of having a cancerous tumor.

105. The method of Claim 102, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

106. A method for treating or preventing a cell proliferative disorder associated with increased expression or activity of a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID

NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising administering to a subject in need of such treatment an effective amount of an antagonist of said protein, thereby effectively treating or preventing said cell proliferative disorder.

107. The method of Claim 106, wherein said cell proliferative disorder is cancer.

108. The method of Claim 106, wherein said antagonist is an anti-TAT polypeptide antibody, TAT binding oligopeptide, TAT binding organic molecule or antisense oligonucleotide.

109. A method of binding an antibody, oligopeptide or organic molecule to a cell that expresses a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising contacting said cell with an antibody, oligopeptide or organic molecule that binds to said protein and allowing the binding of the antibody, oligopeptide or organic molecule to said protein to occur, thereby binding said antibody, oligopeptide or organic molecule to said cell.

110. The method of Claim 109, wherein said antibody is a monoclonal antibody.

111. The method of Claim 109, wherein said antibody is an antibody fragment.

112. The method of Claim 109, wherein said antibody is a chimeric or a humanized antibody.

113. The method of Claim 109, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.

114. The method of Claim 109, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.

115. The method of Claim 114, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

116. The method of Claim 114, wherein the cytotoxic agent is a toxin.

117. The method of Claim 116, wherein the toxin is selected from the group consisting of

maytansinoid and calicheamicin.

118. The method of Claim 116, wherein the toxin is a maytansinoid.

119. The method of Claim 109, wherein said antibody is produced in bacteria.

120. The method of Claim 109, wherein said antibody is produced in CHO cells.

121. The method of Claim 109, wherein said cell is a cancer cell.

122. The method of Claim 121, wherein said cancer cell is further exposed to radiation treatment or a chemotherapeutic agent.

123. The method of Claim 121, wherein said cancer cell is selected from the group consisting of a breast cancer cell, a colorectal cancer cell, a lung cancer cell, an ovarian cancer cell, a central nervous system cancer cell, a liver cancer cell, a bladder cancer cell, a pancreatic cancer cell, a cervical cancer cell, a melanoma cell and a leukemia cell.

124. The method of Claim 123, wherein said protein is more abundantly expressed by said cancer cell as compared to a normal cell of the same tissue origin.

125. The method of Claim 109 which causes the death of said cell.

126. Use of a nucleic acid as claimed in any of Claims 1 to 5 or 30 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

127. Use of a nucleic acid as claimed in any of Claims 1 to 5 or 30 in the preparation of a medicament for treating a tumor.

128. Use of a nucleic acid as claimed in any of Claims 1 to 5 or 30 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

129. Use of an expression vector as claimed in any of Claims 6, 7 or 31 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

130. Use of an expression vector as claimed in any of Claims 6, 7 or 31 in the preparation of medicament for treating a tumor.

131. Use of an expression vector as claimed in any of Claims 6, 7 or 31 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

132. Use of a host cell as claimed in any of Claims 8, 9, 32, or 33 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

133. Use of a host cell as claimed in any of Claims 8, 9, 32 or 33 in the preparation of a medicament for treating a tumor.

134. Use of a host cell as claimed in any of Claims 8, 9, 32 or 33 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

135. Use of a polypeptide as claimed in any of Claims 11 to 14 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

136. Use of a polypeptide as claimed in any of Claims 11 to 14 in the preparation of a medicament for treating a tumor.

137. Use of a polypeptide as claimed in any of Claims 11 to 14 in the preparation of a

medicament for treatment or prevention of a cell proliferative disorder.

138. Use of an antibody as claimed in any of Claims 15 to 29 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

139. Use of an antibody as claimed in any of Claims 15 to 29 in the preparation of a medicament for treating a tumor.

140. Use of an antibody as claimed in any of Claims 15 to 29 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

141. Use of an oligopeptide as claimed in any of Claims 35 to 44 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

142. Use of an oligopeptide as claimed in any of Claims 35 to 44 in the preparation of a medicament for treating a tumor.

143. Use of an oligopeptide as claimed in any of Claims 35 to 44 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

144. Use of a TAT binding organic molecule as claimed in any of Claims 45 to 54 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

145. Use of a TAT binding organic molecule as claimed in any of Claims 45 to 54 in the preparation of a medicament for treating a tumor.

146. Use of a TAT binding organic molecule as claimed in any of Claims 45 to 54 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

147. Use of a composition of matter as claimed in any of Claims 55 or 56 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

148. Use of a composition of matter as claimed in any of Claims 55 or 56 in the preparation of a medicament for treating a tumor.

149. Use of a composition of matter as claimed in any of Claims 55 or 56 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

150. Use of an article of manufacture as claimed in any of Claims 57 or 58 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

151. Use of an article of manufacture as claimed in any of Claims 57 or 58 in the preparation of a medicament for treating a tumor.

152. Use of an article of manufacture as claimed in any of Claims 57 or 58 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

153. A method for inhibiting the growth of a cell, wherein the growth of said cell is at least in part dependent upon a growth potentiating effect of a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising contacting said protein with an antibody, oligopeptide or organic molecule that binds to said protein, there by inhibiting the growth of said cell.

154. The method of Claim 153, wherein said cell is a cancer cell.

155. The method of Claim 153, wherein said protein is expressed by said cell.

156. The method of Claim 153, wherein the binding of said antibody, oligopeptide or organic molecule to said protein antagonizes a cell growth-potentiating activity of said protein.

157. The method of Claim 153, wherein the binding of said antibody, oligopeptide or organic molecule to said protein induces the death of said cell.

158. The method of Claim 153, wherein said antibody is a monoclonal antibody.

159. The method of Claim 153, wherein said antibody is an antibody fragment.

160. The method of Claim 153, wherein said antibody is a chimeric or a humanized antibody.

161. The method of Claim 153, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.

162. The method of Claim 153, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.

163. The method of Claim 162, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

164. The method of Claim 162, wherein the cytotoxic agent is a toxin.

165. The method of Claim 164, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

166. The method of Claim 164, wherein the toxin is a maytansinoid.

167. The method of Claim 153, wherein said antibody is produced in bacteria.

168. The method of Claim 153, wherein said antibody is produced in CHO cells.

169. The method of Claim 153, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

170. A method of therapeutically treating a tumor in a mammal, wherein the growth of said tumor is at least in part dependent upon a growth potentiating effect of a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising contacting said protein with an antibody, oligopeptide or organic molecule that binds to said protein, thereby effectively treating said tumor.

171. The method of Claim 170, wherein said protein is expressed by cells of said tumor.

172. The method of Claim 170, wherein the binding of said antibody, oligopeptide or organic molecule to said protein antagonizes a cell growth-potentiating activity of said protein.

173. The method of Claim 170, wherein said antibody is a monoclonal antibody.

174. The method of Claim 170, wherein said antibody is an antibody fragment.

175. The method of Claim 170, wherein said antibody is a chimeric or a humanized antibody.

176. The method of Claim 170, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.

177. The method of Claim 170, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.

178. The method of Claim 177, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

179. The method of Claim 177, wherein the cytotoxic agent is a toxin.

180. The method of Claim 179, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

181. The method of Claim 179, wherein the toxin is a maytansinoid.

182. The method of Claim 170, wherein said antibody is produced in bacteria.

183. The method of Claim 170, wherein said antibody is produced in CHO cells.

184. The method of Claim 170, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

185. An isolated antibody that binds to a polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

186. An isolated antibody that binds to a polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

187. The antibody of Claim 185 which is a monoclonal antibody.

188. The antibody of Claim 185 which is an antibody fragment.

189. The antibody of Claim 185 which is a chimeric or a humanized antibody.

190. The antibody of Claim 185 which is conjugated to a growth inhibitory agent.

191. The antibody of Claim 185 which is conjugated to a cytotoxic agent.

192. The antibody of Claim 191, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

193. The antibody of Claim 191, wherein the cytotoxic agent is a toxin.

194. The antibody of Claim 193, wherein the toxin is selected from the group consisting of auristatin, maytansinoid and calicheamicin.

195. The antibody of Claim 193, wherein the toxin is a maytansinoid.

196. The antibody of Claim 185 which is produced in bacteria.

197. The antibody of Claim 185 which is produced in CHO cells.

198. The antibody of Claim 185 which induces death of a cell to which it binds.

199. The antibody of Claim 185 which is detectably labeled.

200. The antibody of Claim 193, wherein the toxin is selected from the group consisting of MMAE (mono-methyl auristatin E), MMAF, (auristatin E valeryl benzylhydrazone), and AFP (Auristatin F phenylene diamine).

201. The antibody of claim 193, wherein the toxin is covalently attached to the antibody by a linker.

202. The antibody of claim 201, wherein the linker is selected from the group consisting of maleimidocaproyl (MC), valine-citrulline (val-cit, vc), citrulline (2-amino-5-ureido pentanoic acid), PAB (p-aminobenzylcarbamoyl), Me (N-methyl-valine citrulline), MC(PEG)6-OH (maleimidocaproyl- polyethylene glycol), SPP (N-Succinimidyl 4-(2-pyridylthio) pentanoate), SMCC (N-Succinimidyl 4-(N-maleimidomethyl) cyclohexane-1 carboxylate), and MC-vc-PAB.

203. The antibody of claim 200, wherein the toxin is MMAE.

204. The antibody of claim 200, wherein the toxin is MMAF.

205. The antibody of claim 202, wherein the linker is MC-vc-PAB and the toxin is MMAE or MMAF.

206. The antibody of any one of claims 185-205, wherein the antibody binds the TAT188 polypeptide.

207. The antibody of claim 206, wherein the antibody inhibits proliferation or promotes cell death of a cell expressing TAT188.

208. The antibody of claim 207, wherein the cell is a cancer cell.

209. The antibody of claim 208, where the cancer cell is selected from the group of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

210. An isolated antibody that competes with binding to the epitopes of TAT188 polypeptide bound by an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195).

211. An isolated antibody having the biological activity of an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195), wherein the biological activity is inhibition of cell proliferation or promotion of cell death in a cell expressing TAT188.

212. An isolated antibody comprising in a corresponding complementary determining region (CDR) an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%,

95%, 96%, 97%, 98%, 99% or 100% of the amino acid sequence of at least 1, 2, 3, 4, 5, or 6 of the CDR(s) of the antibody produced by a hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195).

213. The antibody of any one of claims 185-205, wherein the antibody comprises in a corresponding complementary determining region (CDR) an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid sequence of at least 1, 2, 3, 4, 5, or 6 of the CDR(s) of the antibody produced by a hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195).

214. The antibody of any one of claims 185-205, wherein the antibody exhibits the biological activity of an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195), wherein the biological activity is inhibition of cell proliferation or promotion of cell death in a cell expressing TAT188.

215. The antibody of claim 214, wherein the cell is a cancer cell.

216. The antibody of claim 215, wherein the cancer cell is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

217. A method of inhibiting growth of a cell expressing TAT188, the method comprising contacting the cell with an antibody of any one of claims 185-205.

218. The method of claim 217, wherein the cell is a cancer cell.

219. The method of claim 218, wherein the cancer cell is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

220. The method of claim 219, wherein the cancer cell is a mammalian cell.

221. The method of claim 220, wherein the mammalian cell is a human cell.

222. A method of inhibiting growth of a cell expressing TAT188, the method comprising contacting the cell with an antibody of any one of claims 185-205, wherein the antibody comprises in a corresponding complementary determining region (CDR) an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid sequence of at least 1, 2, 3, 4, 5, or 6 of the CDR(s) of the antibody produced by a hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195).

223. A method of detecting the level of TAT188 polypeptide expressed in a test cell relative to a control cell, the method comprising:

(a) contacting the test cell and the control cell with an isolated anti-TAT188 antibody of claim 210;

(b) detecting binding of the antibody; and

(c) determining the relative binding of the antibody to the test and control cell.

224. The method of claim 223, wherein the test cell and control cell are lysed .

225. The method of claim 223, wherein the test cell is in a tissue.

226. The method of claim 225, wherein the tissue is a tumor tissue.

227. The method of claim 226, wherein the tumor tissue is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

228. A method of detecting the level of TAT188 polypeptide or a polypeptide having at least 80% sequence identity to the amino acid sequence shown in Figure 115 (SEQ ID NO:115) in a test cell relative to a control cell,, the method comprising:

(a) contacting the test cell and the control cell with an isolated antibody of any one of claims 185-189, 196-199, and 210-212;

(b) detecting binding of the antibody; and

(c) determining the relative binding of the antibody to the test and control cell.

229. The method of claim 228, wherein the level of TAT188 polypeptide in the test cell is greater than that in the control cell.

230. The method of claim 229, wherein the method diagnoses cancer in a tissue containing or having contained the test cell.

231. A TAT binding interfering RNA (siRNA) which binds to a nucleic acid having at least 80% sequence identity to:
(a) a nucleotide sequence shown in Figure 37 (SEQ ID NO:37; and
(b) the complement of (a), wherein the siRNA reduces expression of TAT188.

232. An expression vector comprising the siRNA of Claim 231.

233. The expression vector of Claim 232, wherein said siRNA is operably linked to control sequences recognized by a host cell transfected with the vector.

234. A host cell comprising the expression vector of Claim 233.

235. A composition of matter comprising:
(a) the antibody of Claim 185, or
(b) the siRNA of claim 231,
in combination with a carrier.

236. The composition of matter of Claim 235, wherein said carrier is a pharmaceutically acceptable carrier.

237. An article of manufacture:
(a) a container; and
(b) the composition of matter of Claim 235 contained within said container.

238. The article of manufacture of Claim 237, further comprising a label affixed to said container, or a package insert included with said container, referring to the use of said composition of matter for the therapeutic treatment of or the diagnostic detection of a cancer.

239. A method of inhibiting the growth of a cancer cell that expresses a polypeptide having at least 80% amino acid sequence identity to the amino acid sequence shown in Figure 115 (SEQ ID NO:115), said method comprising contacting said cancer cell with a siRNA that binds to a nucleic acid encoding the amino acid in said cancer cell, thereby inhibiting the growth of said cancer cell.

240. The method of claim 239, wherein the nucleic acid has the sequence shown in Figure 37 (SEQ ID NO:37).

241. The method of Claim 228, wherein the detecting the level of expression of the polypeptide

comprises employing an antibody in an immunohistochemistry analysis.

242. A method for treating or preventing a cell proliferative disorder associated with increased expression or activity of a polypeptide having at least 80% amino acid sequence identity to the amino acid sequence shown in Figure 115 (SEQ ID NO:115), said method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a TAT188 polypeptide.

243. The method of claim 242, wherein said antagonist is an isolated anti-TAT188 polypeptide antibody of any one of claims 1-21 and 23-28.

244. The method of claim 243, wherein the cell proliferative disorder is cancer.

245. The method of claim 244, wherein the cancer is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

Yet further embodiments of the present invention will be evident to the skilled artisan upon a reading of the present specification.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a nucleotide sequence (SEQ ID NO:1) of a TAT161 cDNA, wherein SEQ ID NO:1 is a clone designated herein as "DNA77507".

Figures 2A-B show a nucleotide sequence (SEQ ID NO:2) of a TAT101 cDNA, wherein SEQ ID NO:2 is a clone designated herein as "DNA80894".

Figure 3 shows a nucleotide sequence (SEQ ID NO:3) of a TAT157 cDNA, wherein SEQ ID NO:3 is a clone designated herein as "DNA82343".

Figure 4 shows a nucleotide sequence (SEQ ID NO:4) of a TAT160 cDNA, wherein SEQ ID NO:4 is a clone designated herein as "DNA87994".

Figure 5 shows a nucleotide sequence (SEQ ID NO:5) of a TAT158 cDNA, wherein SEQ ID NO:5 is a clone designated herein as "DNA88131".

Figure 6 shows a nucleotide sequence (SEQ ID NO:6) of a TAT110 cDNA, wherein SEQ ID NO:6 is a clone designated herein as "DNA95930".

Figure 7 shows a nucleotide sequence (SEQ ID NO:7) of a TAT210 cDNA, wherein SEQ ID NO:7 is a clone designated herein as "DNA95930-1".

Figure 8 shows a nucleotide sequence (SEQ ID NO:8) of a TAT159 cDNA, wherein SEQ ID NO:8 is a clone designated herein as "DNA96917".

Figure 9 shows a nucleotide sequence (SEQ ID NO:9) of a TAT112 cDNA, wherein SEQ ID NO:9 is a clone designated herein as "DNA96930".

Figure 10 shows a nucleotide sequence (SEQ ID NO:10) of a TAT147 cDNA, wherein SEQ ID NO:10 is a clone designated herein as "DNA96936".

Figure 11 shows a nucleotide sequence (SEQ ID NO:11) of a TAT145 cDNA, wherein SEQ ID NO:11 is a clone designated herein as "DNA98565".

Figure 12 shows a nucleotide sequence (SEQ ID NO:12) of a TAT152 cDNA, wherein SEQ ID NO:12 is a clone designated herein as "DNA246435".

Figure 13 shows a nucleotide sequence (SEQ ID NO:13) of a TAT162 cDNA, wherein SEQ ID NO:13 is a clone designated herein as "DNA98591".

Figure 14 shows a nucleotide sequence (SEQ ID NO:14) of a TAT114 cDNA, wherein SEQ ID NO:14 is a clone designated herein as "DNA108809".

Figure 15 shows a nucleotide sequence (SEQ ID NO:15) of a TAT119 cDNA, wherein SEQ ID NO:15 is a clone designated herein as "DNA119488".

Figure 16 shows a nucleotide sequence (SEQ ID NO:16) of a TAT103 cDNA, wherein SEQ ID NO:16 is a clone designated herein as "DNA143493".

Figures 17A-B show a nucleotide sequence (SEQ ID NO:17) of a TAT130 cDNA, wherein SEQ ID NO:17 is a clone designated herein as "DNA167234".

Figure 18 shows a nucleotide sequence (SEQ ID NO:18) of a TAT166 cDNA, wherein SEQ ID NO:18 is a clone designated herein as "DNA235621".

Figure 19 shows a nucleotide sequence (SEQ ID NO:19) of a TAT132 cDNA, wherein SEQ ID NO:19 is a clone designated herein as "DNA176766".

Figure 20 shows a nucleotide sequence (SEQ ID NO:20) of a TAT150 cDNA, wherein SEQ ID NO:20 is a clone designated herein as "DNA236463".

Figure 21 shows a nucleotide sequence (SEQ ID NO:21) of a TAT129 cDNA, wherein SEQ ID NO:21 is a clone designated herein as "DNA181162".

Figure 22 shows a nucleotide sequence (SEQ ID NO:22) of a TAT111 cDNA, wherein SEQ ID NO:22 is a clone designated herein as "DNA188221".

Figure 23 shows a nucleotide sequence (SEQ ID NO:23) of a TAT146 cDNA, wherein SEQ ID NO:23 is a clone designated herein as "DNA233876".

Figure 24 shows a nucleotide sequence (SEQ ID NO:24) of a TAT148 cDNA, wherein SEQ ID NO:24 is a clone designated herein as "DNA193891".

Figure 25 shows a nucleotide sequence (SEQ ID NO:25) of a TAT187 cDNA, wherein SEQ ID NO:25 is a clone designated herein as "DNA248170".

Figure 26 shows a nucleotide sequence (SEQ ID NO:26) of a TAT118 cDNA, wherein SEQ ID NO:26 is a clone designated herein as "DNA194628".

Figure 27 shows a nucleotide sequence (SEQ ID NO:27) of a TAT167 cDNA, wherein SEQ ID NO:27 is a clone designated herein as "DNA246415".

Figure 28 shows a nucleotide sequence (SEQ ID NO:28) of a TAT123 cDNA, wherein SEQ ID

NO:28 is a clone designated herein as "DNA210499".

Figure 29 shows a nucleotide sequence (SEQ ID NO:29) of a TAT211 cDNA, wherein SEQ ID NO:29 is a clone designated herein as "DNA219894".

Figure 30 shows a nucleotide sequence (SEQ ID NO:30) of a TAT113 cDNA, wherein SEQ ID NO:30 is a clone designated herein as "DNA215609".

Figure 31 shows a nucleotide sequence (SEQ ID NO:31) of a TAT128 cDNA, wherein SEQ ID NO:31 is a clone designated herein as "DNA220432".

Figures 32A-B show a nucleotide sequence (SEQ ID NO:32) of a TAT164 cDNA, wherein SEQ ID NO:32 is a clone designated herein as "DNA226094".

Figure 33 shows a nucleotide sequence (SEQ ID NO:33) of a TAT122 cDNA, wherein SEQ ID NO:33 is a clone designated herein as "DNA226165".

Figure 34 shows a nucleotide sequence (SEQ ID NO:34) of a TAT117 cDNA, wherein SEQ ID NO:34 is a clone designated herein as "DNA226237".

Figure 35 shows a nucleotide sequence (SEQ ID NO:35) of a TAT168 cDNA, wherein SEQ ID NO:35 is a clone designated herein as "DNA246450".

Figure 36 shows a nucleotide sequence (SEQ ID NO:36) of a TAT144 cDNA, wherein SEQ ID NO:36 is a clone designated herein as "DNA226456".

Figure 37 shows a nucleotide sequence (SEQ ID NO:37) of a TAT188 cDNA, wherein SEQ ID NO:37 is a clone designated herein as "DNA237637".

Figure 38 shows a nucleotide sequence (SEQ ID NO:38) of a TAT126 cDNA, wherein SEQ ID NO:38 is a clone designated herein as "DNA226539".

Figure 39 shows a nucleotide sequence (SEQ ID NO:39) of a TAT151 cDNA, wherein SEQ ID NO:39 is a clone designated herein as "DNA236511".

Figure 40 shows a nucleotide sequence (SEQ ID NO:40) of a TAT115 cDNA, wherein SEQ ID NO:40 is a clone designated herein as "DNA226771".

Figure 41 shows a nucleotide sequence (SEQ ID NO:41) of a TAT163 cDNA, wherein SEQ ID NO:41 is a clone designated herein as "DNA227087".

Figure 42 shows a nucleotide sequence (SEQ ID NO:42) of a TAT227 cDNA, wherein SEQ ID NO:42 is a clone designated herein as "DNA266307".

Figure 43 shows a nucleotide sequence (SEQ ID NO:43) of a TAT228 cDNA, wherein SEQ ID NO:43 is a clone designated herein as "DNA266311".

Figure 44 shows a nucleotide sequence (SEQ ID NO:44) of a TAT229 cDNA, wherein SEQ ID NO:44 is a clone designated herein as "DNA266312".

Figure 45 shows a nucleotide sequence (SEQ ID NO:45) of a TAT230 cDNA, wherein SEQ ID NO:45 is a clone designated herein as "DNA266313".

Figure 46 shows a nucleotide sequence (SEQ ID NO:46) of a TAT121 cDNA, wherein SEQ ID NO:46 is a clone designated herein as "DNA227224".

Figure 47 shows a nucleotide sequence (SEQ ID NO:47) of a TAT183 cDNA, wherein SEQ ID NO:47 is a clone designated herein as "DNA247486".

Figure 48 shows a nucleotide sequence (SEQ ID NO:48) of a TAT165 cDNA, wherein SEQ ID NO:48 is a clone designated herein as "DNA227578".

Figure 49 shows a nucleotide sequence (SEQ ID NO:49) of a TAT131 cDNA, wherein SEQ ID NO:49 is a clone designated herein as "DNA227800".

Figure 50 shows a nucleotide sequence (SEQ ID NO:50) of a TAT140 cDNA, wherein SEQ ID NO:50 is a clone designated herein as "DNA227904".

Figure 51 shows a nucleotide sequence (SEQ ID NO:51) of a TAT127 cDNA, wherein SEQ ID NO:51 is a clone designated herein as "DNA228199".

Figure 52 shows a nucleotide sequence (SEQ ID NO:52) of a TAT116 cDNA, wherein SEQ ID NO:52 is a clone designated herein as "DNA228201".

Figure 53 shows a nucleotide sequence (SEQ ID NO:53) of a TAT189 cDNA, wherein SEQ ID NO:53 is a clone designated herein as "DNA247488".

Figure 54 shows a nucleotide sequence (SEQ ID NO:54) of a TAT190 cDNA, wherein SEQ ID NO:54 is a clone designated herein as "DNA236538".

Figure 55 shows a nucleotide sequence (SEQ ID NO:55) of a TAT191 cDNA, wherein SEQ ID NO:55 is a clone designated herein as "DNA247489".

Figure 56 shows a nucleotide sequence (SEQ ID NO:56) of a TAT133 cDNA, wherein SEQ ID NO:56 is a clone designated herein as "DNA228211".

Figure 57 shows a nucleotide sequence (SEQ ID NO:57) of a TAT186 cDNA, wherein SEQ ID NO:57 is a clone designated herein as "DNA233937".

Figure 58 shows a nucleotide sequence (SEQ ID NO:58) of a TAT120 cDNA, wherein SEQ ID NO:58 is a clone designated herein as "DNA228993".

Figure 59 shows a nucleotide sequence (SEQ ID NO:59) of a TAT124 cDNA, wherein SEQ ID NO:59 is a clone designated herein as "DNA228994".

Figure 60 shows a nucleotide sequence (SEQ ID NO:60) of a TAT105 cDNA, wherein SEQ ID NO:60 is a clone designated herein as "DNA229410".

Figures 61A-B show a nucleotide sequence (SEQ ID NO:61) of a TAT107 cDNA, wherein SEQ ID NO:61 is a clone designated herein as "DNA229411".

Figures 62A-B show a nucleotide sequence (SEQ ID NO:62) of a TAT108 cDNA, wherein SEQ ID NO:62 is a clone designated herein as "DNA229413".

Figures 63A-B show a nucleotide sequence (SEQ ID NO:63) of a TAT139 cDNA, wherein SEQ ID NO:63 is a clone designated herein as "DNA229700".

Figure 64 shows a nucleotide sequence (SEQ ID NO:64) of a TAT143 cDNA, wherein SEQ ID NO:64 is a clone designated herein as "DNA231312".

Figure 65 shows a nucleotide sequence (SEQ ID NO:65) of a TAT100 cDNA, wherein SEQ ID

NO:65 is a clone designated herein as "DNA231542".

Figure 66 shows a nucleotide sequence (SEQ ID NO:66) of a TAT284 cDNA, wherein SEQ ID NO:66 is a clone designated herein as "DNA231542-1".

Figure 67 shows a nucleotide sequence (SEQ ID NO:67) of a TAT285 cDNA, wherein SEQ ID NO:67 is a clone designated herein as "DNA231542-2".

Figure 68 shows a nucleotide sequence (SEQ ID NO:68) of a TAT285-1 cDNA, wherein SEQ ID NO:68 is a clone designated herein as "DNA297393".

Figure 69 shows a nucleotide sequence (SEQ ID NO:69) of a TAT125 cDNA, wherein SEQ ID NO:69 is a clone designated herein as "DNA232754".

Figure 70 shows a nucleotide sequence (SEQ ID NO:70) of a TAT149 cDNA, wherein SEQ ID NO:70 is a clone designated herein as "DNA234833".

Figure 71 shows a nucleotide sequence (SEQ ID NO:71) of a TAT231 cDNA, wherein SEQ ID NO:71 is a clone designated herein as "DNA268022".

Figure 72 shows a nucleotide sequence (SEQ ID NO:72) of a TAT153 cDNA, wherein SEQ ID NO:72 is a clone designated herein as "DNA236246".

Figure 73 shows a nucleotide sequence (SEQ ID NO:73) of a TAT104 cDNA, wherein SEQ ID NO:73 is a clone designated herein as "DNA236343".

Figure 74 shows a nucleotide sequence (SEQ ID NO:74) of a TAT141 cDNA, wherein SEQ ID NO:74 is a clone designated herein as "DNA236493".

Figure 75 shows a nucleotide sequence (SEQ ID NO:75) of a TAT102 cDNA, wherein SEQ ID NO:75 is a clone designated herein as "DNA236534".

Figure 76 shows a nucleotide sequence (SEQ ID NO:76) of a TAT109 cDNA, wherein SEQ ID NO:76 is a clone designated herein as "DNA246430".

Figure 77 shows a nucleotide sequence (SEQ ID NO:77) of a TAT142 cDNA, wherein SEQ ID NO:77 is a clone designated herein as "DNA247480".

Figures 78A-B show a nucleotide sequence (SEQ ID NO:78) of a TAT106 cDNA, wherein SEQ ID NO:78 is a clone designated herein as "DNA264454".

Figure 79 shows the amino acid sequence (SEQ ID NO:79) derived from the coding sequence of SEQ ID NO:1 shown in Figure 1.

Figure 80 shows the amino acid sequence (SEQ ID NO:80) derived from the coding sequence of SEQ ID NO:2 shown in Figure 2.

Figure 81 shows the amino acid sequence (SEQ ID NO:81) derived from the coding sequence of SEQ ID NO:3 shown in Figure 3.

Figure 82 shows the amino acid sequence (SEQ ID NO:82) derived from the coding sequence of SEQ ID NO:4 shown in Figure 4.

Figure 83 shows the amino acid sequence (SEQ ID NO:83) derived from the coding sequence of SEQ ID NO:5 shown in Figure 5.

Figure 84 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID NO:6 shown in Figure 6.

Figure 85 shows the amino acid sequence (SEQ ID NO:85) derived from the coding sequence of SEQ ID NO:7 shown in Figure 7.

Figure 86 shows the amino acid sequence (SEQ ID NO:86) derived from the coding sequence of SEQ ID NO:8 shown in Figure 8.

Figure 87 shows the amino acid sequence (SEQ ID NO:87) derived from the coding sequence of SEQ ID NO:9 shown in Figure 9.

Figure 88 shows the amino acid sequence (SEQ ID NO:88) derived from the coding sequence of SEQ ID NO:10 shown in Figure 10.

Figure 89 shows the amino acid sequence (SEQ ID NO:89) derived from the coding sequence of SEQ ID NO:11 shown in Figure 11.

Figure 90 shows the amino acid sequence (SEQ ID NO:90) derived from the coding sequence of SEQ ID NO:12 shown in Figure 12.

Figure 91 shows the amino acid sequence (SEQ ID NO:91) derived from the coding sequence of SEQ ID NO:13 shown in Figure 13.

Figure 92 shows the amino acid sequence (SEQ ID NO:92) derived from the coding sequence of SEQ ID NO:14 shown in Figure 14.

Figure 93 shows the amino acid sequence (SEQ ID NO:93) derived from the coding sequence of SEQ ID NO:15 shown in Figure 15.

Figure 94 shows the amino acid sequence (SEQ ID NO:94) derived from the coding sequence of SEQ ID NO:16 shown in Figure 16.

Figure 95 shows the amino acid sequence (SEQ ID NO:95) derived from the coding sequence of SEQ ID NO:17 shown in Figures 17A-B.

Figure 96 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:18 shown in Figure 18.

Figure 97 shows the amino acid sequence (SEQ ID NO:97) derived from the coding sequence of SEQ ID NO:19 shown in Figure 19.

Figure 98 shows the amino acid sequence (SEQ ID NO:98) derived from the coding sequence of SEQ ID NO:20 shown in Figure 20.

Figure 99 shows the amino acid sequence (SEQ ID NO:99) derived from the coding sequence of SEQ ID NO:21 shown in Figure 21.

Figure 100 shows the amino acid sequence (SEQ ID NO:100) derived from the coding sequence of SEQ ID NO:22 shown in Figure 22.

Figure 101 shows the amino acid sequence (SEQ ID NO:101) derived from the coding sequence of SEQ ID NO:23 shown in Figure 23.

Figure 102 shows the amino acid sequence (SEQ ID NO:102) derived from the coding sequence of

SEQ ID NO:24 shown in Figure 24.

Figure 103 shows the amino acid sequence (SEQ ID NO:103) derived from the coding sequence of SEQ ID NO:25 shown in Figure 25.

Figure 104 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:26 shown in Figure 26.

Figure 105 shows the amino acid sequence (SEQ ID NO:105) derived from the coding sequence of SEQ ID NO:27 shown in Figure 27.

Figure 106 shows the amino acid sequence (SEQ ID NO:106) derived from the coding sequence of SEQ ID NO:28 shown in Figure 28.

Figure 107 shows the amino acid sequence (SEQ ID NO:107) derived from the coding sequence of SEQ ID NO:29 shown in Figure 29.

Figure 108 shows the amino acid sequence (SEQ ID NO:108) derived from the coding sequence of SEQ ID NO:30 shown in Figure 30.

Figure 109 shows the amino acid sequence (SEQ ID NO:109) derived from the coding sequence of SEQ ID NO:31 shown in Figure 31.

Figures 110A-B shows the amino acid sequence (SEQ ID NO:110) derived from the coding sequence of SEQ ID NO:32 shown in Figures 32A-B.

Figure 111 shows the amino acid sequence (SEQ ID NO:111) derived from the coding sequence of SEQ ID NO:33 shown in Figure 33.

Figure 112 shows the amino acid sequence (SEQ ID NO:112) derived from the coding sequence of SEQ ID NO:34 shown in Figure 34.

Figure 113 shows the amino acid sequence (SEQ ID NO:113) derived from the coding sequence of SEQ ID NO:35 shown in Figure 35.

Figure 114 shows the amino acid sequence (SEQ ID NO:114) derived from the coding sequence of SEQ ID NO:36 shown in Figure 36.

Figure 115 shows the amino acid sequence (SEQ ID NO:115) derived from the coding sequence of SEQ ID NO:37 shown in Figure 37.

Figure 116 shows the amino acid sequence (SEQ ID NO:116) derived from the coding sequence of SEQ ID NO:38 shown in Figure 38.

Figure 117 shows the amino acid sequence (SEQ ID NO:117) derived from the coding sequence of SEQ ID NO:39 shown in Figure 39.

Figure 118 shows the amino acid sequence (SEQ ID NO:118) derived from the coding sequence of SEQ ID NO:40 shown in Figure 40.

Figure 119 shows the amino acid sequence (SEQ ID NO:119) derived from the coding sequence of SEQ ID NO:41 shown in Figure 41.

Figure 120 shows the amino acid sequence (SEQ ID NO:120) derived from the coding sequence of SEQ ID NO:42 shown in Figure 42.

Figure 121 shows the amino acid sequence (SEQ ID NO:121) derived from the coding sequence of SEQ ID NO:43 shown in Figure 43.

Figure 122 shows the amino acid sequence (SEQ ID NO:122) derived from the coding sequence of SEQ ID NO:44 shown in Figure 44.

Figure 123 shows the amino acid sequence (SEQ ID NO:123) derived from the coding sequence of SEQ ID NO:45 shown in Figure 45.

Figure 124 shows the amino acid sequence (SEQ ID NO:124) derived from the coding sequence of SEQ ID NO:46 shown in Figure 46.

Figure 125 shows the amino acid sequence (SEQ ID NO:125) derived from the coding sequence of SEQ ID NO:47 shown in Figure 47.

Figure 126 shows the amino acid sequence (SEQ ID NO:126) derived from the coding sequence of SEQ ID NO:48 shown in Figure 48.

Figure 127 shows the amino acid sequence (SEQ ID NO:127) derived from the coding sequence of SEQ ID NO:49 shown in Figure 49.

Figure 128 shows the amino acid sequence (SEQ ID NO:128) derived from the coding sequence of SEQ ID NO:50 shown in Figure 50.

Figure 129 shows the amino acid sequence (SEQ ID NO:129) derived from the coding sequence of SEQ ID NO:51 shown in Figure 51.

Figure 130 shows the amino acid sequence (SEQ ID NO:130) derived from the coding sequence of SEQ ID NO:52 shown in Figure 52.

Figure 131 shows the amino acid sequence (SEQ ID NO:131) derived from the coding sequence of SEQ ID NO:53 shown in Figure 53.

Figure 132 shows the amino acid sequence (SEQ ID NO:132) derived from the coding sequence of SEQ ID NO:54 shown in Figure 54.

Figure 133 shows the amino acid sequence (SEQ ID NO:133) derived from the coding sequence of SEQ ID NO:55 shown in Figure 55.

Figure 134 shows the amino acid sequence (SEQ ID NO:134) derived from the coding sequence of SEQ ID NO:56 shown in Figure 56.

Figure 135 shows the amino acid sequence (SEQ ID NO:135) derived from the coding sequence of SEQ ID NO:57 shown in Figure 57.

Figure 136 shows the amino acid sequence (SEQ ID NO:136) derived from the coding sequence of SEQ ID NO:58 shown in Figure 58.

Figure 137 shows the amino acid sequence (SEQ ID NO:137) derived from the coding sequence of SEQ ID NO:59 shown in Figure 59.

Figure 138 shows the amino acid sequence (SEQ ID NO:138) derived from the coding sequence of SEQ ID NO:60 shown in Figure 60.

Figure 139 shows the amino acid sequence (SEQ ID NO:139) derived from the coding sequence of

SEQ ID NO:61 shown in Figures 61A-B.

Figure 140 shows the amino acid sequence (SEQ ID NO:140) derived from the coding sequence of SEQ ID NO:62 shown in Figures 62A-B.

Figure 141 shows the amino acid sequence (SEQ ID NO:141) derived from the coding sequence of SEQ ID NO:63 shown in Figures 63A-B.

Figure 142 shows the amino acid sequence (SEQ ID NO:142) derived from the coding sequence of SEQ ID NO:64 shown in Figure 64.

Figure 143 shows the amino acid sequence (SEQ ID NO:143) derived from the coding sequence of SEQ ID NO:66 shown in Figure 66.

Figure 144 shows the amino acid sequence (SEQ ID NO:144) derived from the coding sequence of SEQ ID NO:67 shown in Figure 67.

Figure 145 shows the amino acid sequence (SEQ ID NO:145) derived from the coding sequence of SEQ ID NO:68 shown in Figure 68.

Figure 146 shows the amino acid sequence (SEQ ID NO:146) derived from the coding sequence of SEQ ID NO:69 shown in Figure 69.

Figure 147 shows the amino acid sequence (SEQ ID NO:147) derived from the coding sequence of SEQ ID NO:70 shown in Figure 70.

Figure 148 shows the amino acid sequence (SEQ ID NO:148) derived from the coding sequence of SEQ ID NO:71 shown in Figure 71.

Figure 149 shows the amino acid sequence (SEQ ID NO:149) derived from the coding sequence of SEQ ID NO:73 shown in Figure 73.

Figure 150 shows the amino acid sequence (SEQ ID NO:150) derived from the coding sequence of SEQ ID NO:74 shown in Figure 74.

Figure 151 shows the amino acid sequence (SEQ ID NO:151) derived from the coding sequence of SEQ ID NO:75 shown in Figure 75.

Figure 152 shows the amino acid sequence (SEQ ID NO:152) derived from the coding sequence of SEQ ID NO:76 shown in Figure 76.

Figure 153 shows the amino acid sequence (SEQ ID NO:153) derived from the coding sequence of SEQ ID NO:77 shown in Figure 77.

Figure 154 shows the amino acid sequence (SEQ ID NO:154) derived from the coding sequence of SEQ ID NO:78 shown in Figures 78A-B.

Figure 155 is a diagram depicting the three dimensional structure of the E16 polypeptide (TAT188) as a subunit of the sodium ion independent large neutral amino acid transporter.

Figure 156 shows the graphical results of FACS plots demonstrating binding of anti-TAT188 (anti-E16) antibodies to PC3 cell surface (green plot) and reduction in binding with reduction in TAT188 expression (red plot).

Figure 157 shows internalization of anti-TAT188 antibody after binding to the surface of TAT188-

expressing PC3 cells.

Figure 158 is a bar graph showing changes in TAT188 amino acid transport activity in the presence of anti-TAT188 antibody with time.

Figure 159 provides plots of cell viability in the presence of increasing amounts of MC-vc-PAB-MMAE toxin-conjugated anti-TAT188 antibody in PC3 and Colo205 cells.

Figure 160 is a plot of cell viability in the presence of increasing amounts of MC-vc-PAB-MMAE or MC-vc-PAB-MMAF toxin-conjugated anti-TAT188 antibody in Colo205 cells.

Figure 161shows the results anti-human E16 antibody binding cells endogenously expressing E16 polypeptide, where the cells are monkey COS7 cells, human MCF10A breast cancer cells, and mouse NIH-3T3 cells. Figure 161A shows the FACS results of binding of anti-E16 antibodies 3B5, 12B9, and 12G12 to monkey COS. Figure 161B shows binding of anti-E16 3B5 to human breast tumor cell line MCF10A and no binding to mouse NIH-3T3 cells.

Figure 162 is a plot of mean tumor volume changes over time in xenograft mice administered naked (not toxin-conjugated) anti-E16 antibodies.

Figure 163 is a graph of mean tumor volume changes over time in xenograft mice administerd toxin-conjugated anti-E16 antibodies, where the toxin was either MC-vc-PAB-MMAE or MC-vc-PAB-MMAF.

Figure 164 shows that E16-GFP fusion polypeptide expression is reduced in cells transfected with siRNA that targets the E16 gene. The top panel are photographs of the results of bioluminescence assays for the presence of GFP. The lower panel are photographs of phase contrast visualization showing the GFP bioluminescence reduction was not caused by reduction in cell number.

Figure 165 is a Western blot showing that E16-GFP fusion polypeptide expression is inhibited in PC3 cells transiently transfected with E16 siRNA. β-tubulin is a gel loading control.

Figure 166 shows that transient transfection of PC3 cells with E16 siRNA is associated with a reduction in amino acid transport function, consistent with reduced expression of E16.

Figure 167 is a plot of cell proliferation over time with and without transient transfection with E16 siRNA. Cell proliferation is reduced in PC3 cells transiently transfected with E16.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. Definitions

The terms "TAT polypeptide" and "TAT" as used herein and when immediately followed by a numerical designation, refer to various polypeptides, wherein the complete designation (i.e.,TAT/number) refers to specific polypeptide sequences as described herein. The terms "TAT/number polypeptide" and "TAT/number" wherein the term "number" is provided as an actual numerical designation as used herein encompass native sequence polypeptides, polypeptide variants and fragments of native sequence polypeptides and polypeptide variants (which are further defined herein). The TAT polypeptides described herein may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods. The term "TAT polypeptide" refers to each individual TAT/number

polypeptide disclosed herein. All disclosures in this specification which refer to the "TAT polypeptide" refer to each of the polypeptides individually as well as jointly. For example, descriptions of the preparation of, purification of, derivation of, formation of antibodies to or against, formation of TAT binding oligopeptides to or against, formation of TAT binding organic molecules to or against, administration of, compositions containing, treatment of a disease with, etc., pertain to each polypeptide of the invention individually. The term "TAT polypeptide" also includes variants of the TAT/number polypeptides disclosed herein.

A "native sequence TAT polypeptide" comprises a polypeptide having the same amino acid sequence as the corresponding TAT polypeptide derived from nature. Such native sequence TAT polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence TAT polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of the specific TAT polypeptide (*e.g.*, an extracellular domain sequence), naturally-occurring variant forms (*e.g.*, alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In certain embodiments of the invention, the native sequence TAT polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acids sequences shown in the accompanying figures. Start and stop codons (if indicated) are shown in bold font and underlined in the figures. Nucleic acid residues indicated as "N" in the accompanying figures are any nucleic acid residue. However, while the TAT polypeptides disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino acid position 1 in the figures, it is conceivable and possible that other methionine residues located either upstream or downstream from the amino acid position 1 in the figures may be employed as the starting amino acid residue for the TAT polypeptides.

The TAT polypeptide "extracellular domain" or "ECD" refers to a form of the TAT polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a TAT polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domains identified for the TAT polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified herein. Optionally, therefore, an extracellular domain of a TAT polypeptide may contain from about 5 or fewer amino acids on either side of the transmembrane domain/extracellular domain boundary as identified in the Examples or specification and such polypeptides, with or without the associated signal peptide, and nucleic acid encoding them, are contemplated by the present invention.

The approximate location of the "signal peptides" of the various TAT polypeptides disclosed herein may be shown in the present specification and/or the accompanying figures. It is noted, however, that the C-terminal boundary of a signal peptide may vary, but most likely by no more than about 5 amino acids on either side of the signal peptide C-terminal boundary as initially identified herein, wherein the C-terminal boundary of the signal peptide may be identified pursuant to criteria routinely employed in the art for identifying that type of amino acid sequence element (e.g., Nielsen et al., Prot. Eng. 10:1-6 (1997) and von Heinje et al.,

Nucl. Acids. Res. 14:4683-4690 (1986)). Moreover, it is also recognized that, in some cases, cleavage of a signal sequence from a secreted polypeptide is not entirely uniform, resulting in more than one secreted species. These mature polypeptides, where the signal peptide is cleaved within no more than about 5 amino acids on either side of the C-terminal boundary of the signal peptide as identified herein, and the polynucleotides encoding them, are contemplated by the present invention.

"TAT polypeptide variant" means a TAT polypeptide, preferably an active TAT polypeptide, as defined herein having at least about 80% amino acid sequence identity with a full-length native sequence TAT polypeptide sequence as disclosed herein, a TAT polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a TAT polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length TAT polypeptide sequence as disclosed herein (such as those encoded by a nucleic acid that represents only a portion of the complete coding sequence for a full-length TAT polypeptide). Such TAT polypeptide variants include, for instance, TAT polypeptides wherein one or more amino acid residues are added, or deleted, at the N– or C-terminus of the full-length native amino acid sequence. Ordinarily, a TAT polypeptide variant will have at least about 80% amino acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% amino acid sequence identity, to a full-length native sequence TAT polypeptide sequence as disclosed herein, a TAT polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a TAT polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length TAT polypeptide sequence as disclosed herein. Ordinarily, TAT variant polypeptides are at least about 10 amino acids in length, alternatively at least about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600 amino acids in length, or more. Optionally, TAT variant polypeptides will have no more than one conservative amino acid substitution as compared to the native TAT polypeptide sequence, alternatively no more than 2, 3, 4, 5, 6, 7, 8, 9, or 10 conservative amino acid substitution as compared to the native TAT polypeptide sequence.

"Percent (%) amino acid sequence identity" with respect to the TAT polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific TAT polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is

provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

100 times the fraction X/Y

where X is the number of amino acid residues scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations using this method, Tables 2 and 3 demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "TAT", wherein "TAT" represents the amino acid sequence of a hypothetical TAT polypeptide of interest, "Comparison Protein" represents the amino acid sequence of a polypeptide against which the "TAT" polypeptide of interest is being compared, and "X, "Y" and "Z" each represent different hypothetical amino acid residues. Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program.

"TAT variant polynucleotide" or "TAT variant nucleic acid sequence" means a nucleic acid molecule which encodes a TAT polypeptide, preferably an active TAT polypeptide, as defined herein and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length native sequence TAT polypeptide sequence as disclosed herein, a full-length native sequence TAT polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a TAT polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length TAT polypeptide sequence as disclosed herein (such as those encoded by a nucleic acid that represents only a portion of the complete coding sequence for a full-length TAT polypeptide). Ordinarily, a TAT variant polynucleotide will have at least about 80% nucleic acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% nucleic acid sequence identity with a nucleic acid sequence encoding a full-length native sequence TAT

polypeptide sequence as disclosed herein, a full-length native sequence TAT polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a TAT polypeptide, with or without the signal sequence, as disclosed herein or any other fragment of a full-length TAT polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

Ordinarily, TAT variant polynucleotides are at least about 5 nucleotides in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, or 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length.

"Percent (%) nucleic acid sequence identity" with respect to TAT-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the TAT nucleic acid sequence of interest, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. For purposes herein, however, % nucleic acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for nucleic acid sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

100 times the fraction W/Z

where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will

be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 4 and 5, demonstrate how to calculate the % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "TAT-DNA", wherein "TAT-DNA" represents a hypothetical TAT-encoding nucleic acid sequence of interest, "Comparison DNA" represents the nucleotide sequence of a nucleic acid molecule against which the "TAT-DNA" nucleic acid molecule of interest is being compared, and "N", "L" and "V" each represent different hypothetical nucleotides. Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program.

In other embodiments, TAT variant polynucleotides are nucleic acid molecules that encode a TAT polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding a full-length TAT polypeptide as disclosed herein. TAT variant polypeptides may be those that are encoded by a TAT variant polynucleotide.

The term "full-length coding region" when used in reference to a nucleic acid encoding a TAT polypeptide refers to the sequence of nucleotides which encode the full-length TAT polypeptide of the invention (which is often shown between start and stop codons, inclusive thereof, in the accompanying figures). The term "full-length coding region" when used in reference to an ATCC deposited nucleic acid refers to the TAT polypeptide-encoding portion of the cDNA that is inserted into the vector deposited with the ATCC (which is often shown between start and stop codons, inclusive thereof, in the accompanying figures).

"Isolated," when used to describe the various TAT polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the TAT polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" TAT polypeptide-encoding nucleic acid or other polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the polypeptide-encoding nucleic acid. An isolated polypeptide-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated polypeptide-encoding nucleic acid molecules therefore are distinguished from the specific polypeptide-encoding nucleic acid molecule as it exists in natural cells. However, an isolated polypeptide-

encoding nucleic acid molecule includes polypeptide-encoding nucleic acid molecules contained in cells that ordinarily express the polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

"Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) overnight hybridization in a solution that employs 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with a 10 minute wash at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) followed by a 10 minute high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

"Moderately stringent conditions" may be identified as described by Sambrook et al., Molecular

Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent that those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a TAT polypeptide or anti-TAT antibody fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

"Active" or "activity" for the purposes herein refers to form(s) of a TAT polypeptide which retain a biological and/or an immunological activity of native or naturally-occurring TAT, wherein "biological" activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring TAT other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring TAT and an "immunological" activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring TAT.

The term "antagonist" is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native TAT polypeptide disclosed herein. In a similar manner, the term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native TAT polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically include agonist or antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native TAT polypeptides, peptides, antisense oligonucleotides, small organic molecules, etc. Methods for identifying agonists or antagonists of a TAT polypeptide may comprise contacting a TAT polypeptide with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the TAT polypeptide.

"Treating" or "treatment" or "alleviation" refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented. A subject or mammal is successfully "treated" for a TAT polypeptide-expressing cancer if, after receiving a therapeutic amount of an anti-TAT antibody, TAT binding oligopeptide or TAT binding organic molecule according to the methods of the present invention, the patient shows observable and/or measurable reduction in or absence of one or more

of the following: reduction in the number of cancer cells or absence of the cancer cells; reduction in the tumor size; inhibition (i.e., slow to some extent and preferably stop) of cancer cell infiltration into peripheral organs including the spread of cancer into soft tissue and bone; inhibition (i.e., slow to some extent and preferably stop) of tumor metastasis; inhibition, to some extent, of tumor growth; and/or relief to some extent, one or more of the symptoms associated with the specific cancer; reduced morbidity and mortality, and improvement in quality of life issues. To the extent the anti-TAT antibody or TAT binding oligopeptide may prevent growth and/or kill existing cancer cells, it may be cytostatic and/or cytotoxic. Reduction of these signs or symptoms may also be felt by the patient.

The above parameters for assessing successful treatment and improvement in the disease are readily measurable by routine procedures familiar to a physician. For cancer therapy, efficacy can be measured, for example, by assessing the time to disease progression (TTP) and/or determining the response rate (RR). Metastasis can be determined by staging tests and by bone scan and tests for calcium level and other enzymes to determine spread to the bone. CT scans can also be done to look for spread to the pelvis and lymph nodes in the area. Chest X-rays and measurement of liver enzyme levels by known methods are used to look for metastasis to the lungs and liver, respectively. Other routine methods for monitoring the disease include transrectal ultrasonography (TRUS) and transrectal needle biopsy (TRNB).

For bladder cancer, which is a more localized cancer, methods to determine progress of disease include urinary cytologic evaluation by cystoscopy, monitoring for presence of blood in the urine, visualization of the urothelial tract by sonography or an intravenous pyelogram, computed tomography (CT) and magnetic resonance imaging (MRI). The presence of distant metastases can be assessed by CT of the abdomen, chest x-rays, or radionuclide imaging of the skeleton.

"Chronic" administration refers to administration of the agent(s) in a continuous mode as opposed to an acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. "Intermittent" administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

"Mammal" for purposes of the treatment of, alleviating the symptoms of or diagnosis of a cancer refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, etc. Preferably, the mammal is human.

Administration "in combination with" one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

"Carriers" as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such

as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN®, polyethylene glycol (PEG), and PLURONICS®.

By "solid phase" or "solid support" is meant a non-aqueous matrix to which an antibody, TAT binding oligopeptide or TAT binding organic molecule of the present invention can adhere or attach. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug (such as a TAT polypeptide, an antibody thereto or a TAT binding oligopeptide) to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

A "small" molecule or "small" organic molecule is defined herein to have a molecular weight below about 500 Daltons.

An "effective amount" of a polypeptide, antibody, TAT binding oligopeptide, TAT binding organic molecule or an agonist or antagonist thereof as disclosed herein is an amount sufficient to carry out a specifically stated purpose. An "effective amount" may be determined empirically and in a routine manner, in relation to the stated purpose.

The term "therapeutically effective amount" refers to an amount of an antibody, polypeptide, TAT binding oligopeptide, TAT binding organic molecule or other drug effective to "treat" a disease or disorder in a subject or mammal. In the case of cancer, the therapeutically effective amount of the drug may reduce the number of cancer cells; reduce the tumor size; inhibit (i.e., slow to some extent and preferably stop) cancer cell infiltration into peripheral organs; inhibit (i.e., slow to some extent and preferably stop) tumor metastasis; inhibit, to some extent, tumor growth; and/or relieve to some extent one or more of the symptoms associated with the cancer. See the definition herein of "treating". To the extent the drug may prevent growth and/or kill existing cancer cells, it may be cytostatic and/or cytotoxic.

A "growth inhibitory amount" of an anti-TAT antibody, TAT polypeptide, TAT binding oligopeptide, TAT siRNA (such as a TAT188 siRNA) or TAT binding organic molecule is an amount capable of inhibiting the growth of a cell, especially tumor, e.g., cancer cell, either *in vitro* or *in vivo*. A "growth inhibitory amount" of an anti-TAT antibody, TAT polypeptide, TAT binding oligopeptide, TAT siRNA (such as a TAT188 siRNA) or TAT binding organic molecule for purposes of inhibiting neoplastic cell growth may be determined empirically and in a routine manner.

A "cytotoxic amount" of an anti-TAT antibody, TAT polypeptide, TAT binding oligopeptid, TAT

siRNA (such as a TAT188 siRNA) e or TAT binding organic molecule is an amount capable of causing the destruction of a cell, especially tumor, e.g., cancer cell, either *in vitro* or *in vivo*. A "cytotoxic amount" of an anti-TAT antibody, TAT polypeptide, TAT binding oligopeptide, TAT siRNA (such as a TAT188 siRNA) or TAT binding organic molecule for purposes of inhibiting neoplastic cell growth may be determined empirically and in a routine manner.

The term "antibody" is used in the broadest sense and specifically covers, for example, single anti-TAT monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies), anti-TAT antibody compositions with polyepitopic specificity, polyclonal antibodies, single chain anti-TAT antibodies, and fragments of anti-TAT antibodies (see below) as long as they exhibit the desired biological or immunological activity. The term "immunoglobulin" (Ig) is used interchangeable with antibody herein.

An "isolated antibody" is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody in situ within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

The basic 4-chain antibody unit is a heterotetrameric glycoprotein composed of two identical light (L) chains and two identical heavy (H) chains (an IgM antibody consists of 5 of the basic heterotetramer unit along with an additional polypeptide called J chain, and therefore contain 10 antigen binding sites, while secreted IgA antibodies can polymerize to form polyvalent assemblages comprising 2-5 of the basic 4-chain units along with J chain). In the case of IgGs, the 4-chain unit is generally about 150,000 daltons. Each L chain is linked to a H chain by one covalent disulfide bond, while the two H chains are linked to each other by one or more disulfide bonds depending on the H chain isotype. Each H and L chain also has regularly spaced intrachain disulfide bridges. Each H chain has at the N-terminus, a variable domain (V_H) followed by three constant domains (C_H) for each of the α and γ chains and four C_H domains for μ and ϵ isotypes. Each L chain has at the N-terminus, a variable domain (V_L) followed by a constant domain (C_L) at its other end. The V_L is aligned with the V_H and the C_L is aligned with the first constant domain of the heavy chain (C_H1). Particular amino acid residues are believed to form an interface between the light chain and heavy chain variable domains. The pairing of a V_H and V_L together forms a single antigen-binding site. For the structure and properties of the different classes of antibodies, see, e.g., Basic and Clinical Immunology, 8th edition, Daniel P. Stites, Abba I. Terr and Tristram G. Parslow (eds.), Appleton & Lange, Norwalk, CT, 1994, page 71 and Chapter 6.

The L chain from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains. Depending on the amino acid sequence of the constant domain of their heavy chains (C_H), immunoglobulins can be assigned to different classes or isotypes. There are five classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, having heavy chains designated α, δ, ϵ, γ, and μ, respectively. The γ and α classes are further divided into subclasses on the basis of relatively minor differences in C_H sequence and function, e.g., humans express the following subclasses: IgG1, IgG2, IgG3, IgG4, IgA1, and IgA2.

The term "variable" refers to the fact that certain segments of the variable domains differ extensively in sequence among antibodies. The V domain mediates antigen binding and define specificity of a particular antibody for its particular antigen. However, the variability is not evenly distributed across the 110-amino acid span of the variable domains. Instead, the V regions consist of relatively invariant stretches called framework regions (FRs) of 15-30 amino acids separated by shorter regions of extreme variability called "hypervariable regions" that are each 9-12 amino acids long. The variable domains of native heavy and light chains each comprise four FRs, largely adopting a β-sheet configuration, connected by three hypervariable regions, which form loops connecting, and in some cases forming part of, the β-sheet structure. The hypervariable regions in each chain are held together in close proximity by the FRs and, with the hypervariable regions from the other chain, contribute to the formation of the antigen-binding site of antibodies (see Kabat et al., Sequences of Proteins of Immunological Interest, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD. (1991)). The constant domains are not involved directly in binding an antibody to an antigen, but exhibit various effector functions, such as participation of the antibody in antibody dependent cellular cytotoxicity (ADCC).

The term "hypervariable region" when used herein refers to the amino acid residues of an antibody which are responsible for antigen-binding. The hypervariable region (HVR) generally comprises amino acid residues from a "complementarity determining region" or "CDR" (e.g. around about residues 24-34 (L1), 50-56 (L2) and 89-97 (L3) in the V_L, and around about 1-35 (H1), 50-65 (H2) and 95-102 (H3) in the V_H; Kabat et al., Sequences of Proteins of Immunological Interest, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD. (1991)) and/or those residues from a "hypervariable loop" (*e.g.* residues 26-32 (L1), 50-52 (L2) and 91-96 (L3) in the V_L, and 26-32 (H1), 53-55 (H2) and 96-101 (H3) in the V_H; Chothia and Lesk J. Mol. Biol. 196:901-917 (1987)). The HVRs are also defined as encompassing the sequences as disclosed in US Application Serial No. , filed February 19, 2004, incorporated herein by reference in its entirety, wherein extended HVRs include some of the amino acid sequence positions within the variable regions that are defined as being in contact with a bound ligand of the antibody based on an analysis of complex crystal structures. As used herein the terms "HVR" and "CDR" are used interchangeably to encompass the extend HVRs defined as follows.

The term "hypervariable region" when used herein refers to the regions of an antibody variable domain which are hypervariable in sequence and/or form structurally defined loops. Generally, antibodies comprise six hypervariable regions; three in the VH (H1, H2, H3), and three in the VL (L1, L2, L3). A number of hypervariable region delineations are in use and are encompassed herein. The Kabat Complementarity Determining Regions (CDRs) are based on sequence variability and are the most commonly used (Kabat *et al., Sequences of Proteins of Immunological Interest*, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD. (1991)). Chothia refers instead to the location of the structural loops (Chothia and Lesk *J. Mol. Biol.* 196:901-917 (1987)). The AbM hypervariable regions represent a compromise between the Kabat CDRs and Chothia structural loops, and are used by Oxford Molecular's AbM antibody modeling software. The "contact" hypervariable regions are based on an analysis of the available complex crystal structures. The residues from each of these hypervariable regions are noted below.

Table HVR

Loop	Kabat	AbM	Chothia	Contact
L1	L24-L34	L24-L34	L26-L32	L30-L36
L2	L50-L56	L50-L56	L50-L52	L46-L55
L3	L89-L97	L89-L97	L91-L96	L89-L96
H1 (Kabat Numbering)	H31-H35B	H26-H35B	H26-H32..34	H30-H35B
H1 (Chothia Numbering)	H31-H35	H26-H35	H26-H32	H30-H35
H2	H50-H65	H50-H58	H53-H55	H47-H58
H3	H95-H102	H95-H102	H96-H101	H93-H101

Hypervariable regions may comprise "extended hypervariable regions" as follows: 24-34 (L1), 50-56 (L2) and 89-97 (L3) in the VL and 26-35 (H1), 50-65 or 49-65 (H2) and 93-102, 94-102 or 95-102 (H3) in the VH. The variable domain residues are numbered according to Kabat *et al., supra* for each of these definitions.

"Framework" or "FR" residues are those variable domain residues other than the hypervariable region residues as herein defined.

The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Furthermore, in contrast to polyclonal antibody preparations which include different antibodies directed against different determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen. In addition to their specificity, the monoclonal antibodies are advantageous in that they may be synthesized uncontaminated by other antibodies. The modifier "monoclonal" is not to be construed as requiring production of the antibody by any particular method. For example, the monoclonal antibodies useful in the present invention may be prepared by the hybridoma methodology first described by Kohler et al., Nature, 256:495 (1975), or may be made using recombinant DNA methods in bacterial, eukaryotic animal or plant cells (see, e.g., U.S. Patent No. 4,816,567). The "monoclonal antibodies" may also be isolated from phage antibody libraries using the techniques described in Clackson et al., Nature, 352:624-628 (1991) and Marks et al., J. Mol. Biol., 222:581-597 (1991), for example.

The monoclonal antibodies herein include "chimeric" antibodies in which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, so long as they exhibit the desired biological activity (see U.S. Patent No. 4,816,567; and Morrison et al., Proc. Natl. Acad. Sci. USA, 81:6851-6855 (1984)). Chimeric antibodies of interest herein include "primatized" antibodies comprising variable domain antigen-binding sequences derived from a non-human primate (*e.g.* Old World Monkey, Ape etc), and human constant region sequences.

An "intact" antibody is one which comprises an antigen-binding site as well as a C_L and at least heavy chain constant domains, C_H1, C_H2 and C_H3. The constant domains may be native sequence constant domains (*e.g.* human native sequence constant domains) or amino acid sequence variant thereof. Preferably, the intact antibody has one or more effector functions.

"Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', $F(ab')_2$, and Fv fragments; diabodies; linear antibodies (see U.S. Patent No. 5,641,870, Example 2; Zapata et al., Protein Eng. 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. The Fab fragment consists of an entire L chain along with the variable region domain of the H chain (V_H), and the first constant domain of one heavy chain (C_H1). Each Fab fragment is monovalent with respect to antigen binding, i.e., it has a single antigen-binding site. Pepsin treatment of an antibody yields a single large $F(ab')_2$ fragment

which roughly corresponds to two disulfide linked Fab fragments having divalent antigen-binding activity and is still capable of cross-linking antigen. Fab' fragments differ from Fab fragments by having additional few residues at the carboxy terminus of the C_H1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. $F(ab')_2$ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

The Fc fragment comprises the carboxy-terminal portions of both H chains held together by disulfides. The effector functions of antibodies are determined by sequences in the Fc region, which region is also the part recognized by Fc receptors (FcR) found on certain types of cells.

"Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This fragment consists of a dimer of one heavy- and one light-chain variable region domain in tight, non-covalent association. From the folding of these two domains emanate six hypervariable loops (3 loops each from the H and L chain) that contribute the amino acid residues for antigen binding and confer antigen binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

"Single-chain Fv" also abbreviated as "sFv" or "scFv" are antibody fragments that comprise the V_H and V_L antibody domains connected into a single polypeptide chain. Preferably, the sFv polypeptide further comprises a polypeptide linker between the V_H and V_L domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenburg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994); Borrebaeck 1995, infra.

The term "diabodies" refers to small antibody fragments prepared by constructing sFv fragments (see preceding paragraph) with short linkers (about 5-10 residues) between the V_H and V_L domains such that inter-chain but not intra-chain pairing of the V domains is achieved, resulting in a bivalent fragment, i.e., fragment having two antigen-binding sites. Bispecific diabodies are heterodimers of two "crossover" sFv fragments in which the V_H and V_L domains of the two antibodies are present on different polypeptide chains. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

"Humanized" forms of non-human (*e.g.*, rodent) antibodies are chimeric antibodies that contain minimal sequence derived from the non-human antibody. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which residues from a hypervariable region of the recipient are replaced by residues from a hypervariable region of a non-human species (donor antibody) such as mouse, rat, rabbit or non-human primate having the desired antibody specificity, affinity, and capability. In some instances, framework region (FR) residues of the human immunoglobulin are replaced by corresponding non-human residues. Furthermore, humanized antibodies may comprise residues that are not found in the recipient antibody or in the donor antibody. These modifications are made to further refine antibody performance. In

general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the hypervariable loops correspond to those of a non-human immunoglobulin and all or substantially all of the FRs are those of a human immunoglobulin sequence. The humanized antibody optionally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. For further details, see Jones et al., Nature 321:522-525 (1986); Riechmann et al., Nature 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol. 2:593-596 (1992).

A "species-dependent antibody," e.g., a mammalian anti-human IgE antibody, is an antibody which has a stronger binding affinity for an antigen from a first mammalian species than it has for a homologue of that antigen from a second mammalian species. Normally, the species-dependent antibody "bind specifically" to a human antigen (i.e., has a binding affinity (Kd) value of no more than about 1×10^{-7} M, preferably no more than about 1×10^{-8} and most preferably no more than about 1×10^{-9} M) but has a binding affinity for a homologue of the antigen from a second non-human mammalian species which is at least about 50 fold, or at least about 500 fold, or at least about 1000 fold, weaker than its binding affinity for the human antigen. The species-dependent antibody can be of any of the various types of antibodies as defined above, but preferably is a humanized or human antibody.

A "TAT binding oligopeptide" is an oligopeptide that binds, preferably specifically, to a TAT polypeptide as described herein. TAT binding oligopeptides may be chemically synthesized using known oligopeptide synthesis methodology or may be prepared and purified using recombinant technology. TAT binding oligopeptides are usually at least about 5 amino acids in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 amino acids in length or more, wherein such oligopeptides that are capable of binding, preferably specifically, to a TAT polypeptide as described herein. TAT binding oligopeptides may be identified without undue experimentation using well known techniques. In this regard, it is noted that techniques for screening oligopeptide libraries for oligopeptides that are capable of specifically binding to a polypeptide target are well known in the art (see, e.g., U.S. Patent Nos. 5,556,762, 5,750,373, 4,708,871, 4,833,092, 5,223,409, 5,403,484, 5,571,689, 5,663,143; PCT Publication Nos. WO 84/03506 and WO84/03564; Geysen et al., Proc. Natl. Acad. Sci. U.S.A., 81:3998-4002 (1984); Geysen et al., Proc. Natl. Acad. Sci. U.S.A., 82:178-182 (1985); Geysen et al., in Synthetic Peptides as Antigens, 130-149 (1986); Geysen et al., J. Immunol. Meth., 102:259-274 (1987); Schoofs et al., J. Immunol., 140:611-616 (1988), Cwirla, S. E. et al. (1990) Proc. Natl. Acad. Sci. USA, 87:6378; Lowman, H.B. et al. (1991) Biochemistry, 30:10832; Clackson, T. et al. (1991) Nature, 352: 624; Marks, J. D. et al. (1991), J. Mol. Biol., 222:581; Kang, A.S. et al. (1991) Proc. Natl. Acad. Sci. USA, 88:8363, and Smith, G. P. (1991) Current Opin. Biotechnol., 2:668).

A "TAT binding organic molecule" is an organic molecule other than an oligopeptide or antibody as defined herein that binds, preferably specifically, to a TAT polypeptide as described herein. TAT binding

organic molecules may be identified and chemically synthesized using known methodology (see, e.g., PCT Publication Nos. WO00/00823 and WO00/39585). TAT binding organic molecules are usually less than about 2000 daltons in size, alternatively less than about 1500, 750, 500, 250 or 200 daltons in size, wherein such organic molecules that are capable of binding, preferably specifically, to a TAT polypeptide as described herein may be identified without undue experimentation using well known techniques. In this regard, it is noted that techniques for screening organic molecule libraries for molecules that are capable of binding to a polypeptide target are well known in the art (see, e.g., PCT Publication Nos. WO00/00823 and WO00/39585).

An "interfering RNA" or "small interfering RNA (siRNA)" is a double stranded RNA molecule less than 30 nucleotides in length that reduces expression of a target gene. A "TAT interfering RNA" or "TAT siRNA" binds, preferably specifically, to a TAT nucleic acid and reduces its expression. This means the expression of the TAT molecule is lower with the interfering RNA present as compared to expression of the TAT molecule in the control where the interfering RNA is not present. TAT interfering RNAs may be identified and synthesized using known methods (Shi Y.*Trends in Genetics* 19(1):9-12 (2003), WO/2003056012 and WO2003064621).

An antibody, oligopeptide, siRNA or other organic molecule "which binds" an antigen of interest, e.g. a tumor-associated polypeptide antigen target, is one that binds the antigen with sufficient affinity such that the antibody, oligopeptide or other organic molecule is useful as a diagnostic and/or therapeutic agent in targeting a cell or tissue expressing the antigen, and does not significantly cross-react with other proteins. In such embodiments, the extent of binding of the antibody, oligopeptide, siRNA or other organic molecule to a "non-target" protein will be less than about 10% of the binding of the antibody, oligopeptide, siRNA or other organic molecule to its particular target protein as determined by fluorescence activated cell sorting (FACS) analysis or radioimmunoprecipitation (RIA). With regard to the binding of an antibody, oligopeptide or other organic molecule to a target molecule, the term "specific binding" or "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide target means binding that is measurably different from a non-specific interaction. Specific binding can be measured, for example, by determining binding of a molecule compared to binding of a control molecule, which generally is a molecule of similar structure that does not have binding activity. For example, specific binding can be determined by competition with a control molecule that is similar to the target, for example, an excess of non-labeled target. In this case, specific binding is indicated if the binding of the labeled target to a probe is competitively inhibited by excess unlabeled target. The term "specific binding" or "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide target as used herein can be exhibited, for example, by a molecule having a Kd for the target of at least about 10^{-4} M, alternatively at least about 10^{-5} M, alternatively at least about 10^{-6} M, alternatively at least about 10^{-7} M, alternatively at least about 10^{-8} M, alternatively at least about 10^{-9} M, alternatively at least about 10^{-10} M, alternatively at least about 10^{-11} M, alternatively at least about 10^{-12} M, or greater. In one embodiment, the term "specific binding" refers to binding where a molecule

binds to a particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

An antibody, oligopeptide, siRNA or other organic molecule that "inhibits the growth of tumor cells expressing a TAT polypeptide" or a "growth inhibitory" antibody, oligopeptide or other organic molecule is one which results in measurable growth inhibition of cancer cells expressing or overexpressing the appropriate TAT polypeptide. The TAT polypeptide may be a transmembrane polypeptide expressed on the surface of a cancer cell or may be a polypeptide that is produced and secreted by a cancer cell. Preferred growth inhibitory anti-TAT antibodies, oligopeptides or organic molecules inhibit growth of TAT-expressing tumor cells by greater than 20%, preferably from about 20% to about 50%, and even more preferably, by greater than 50% (e.g., from about 50% to about 100%) as compared to the appropriate control, the control typically being tumor cells not treated with the antibody, oligopeptide or other organic molecule being tested. In one embodiment, growth inhibition can be measured at an antibody concentration of about 0.1 to 30 μg/ml or about 0.5 nM to 200 nM in cell culture, where the growth inhibition is determined 1-10 days after exposure of the tumor cells to the antibody. Growth inhibition of tumor cells *in vivo* can be determined in various ways such as is described in the Experimental Examples section below. The antibody is growth inhibitory in vivo if administration of the anti-TAT antibody at about 1 μg/kg to about 100 mg/kg body weight results in reduction in tumor size or tumor cell proliferation within about 5 days to 3 months from the first administration of the antibody, preferably within about 5 to 30 days.

An antibody, oligopeptide, siRNA or other organic molecule which "induces apoptosis" is one which induces programmed cell death as determined by binding of annexin V, fragmentation of DNA, cell shrinkage, dilation of endoplasmic reticulum, cell fragmentation, and/or formation of membrane vesicles (called apoptotic bodies). The cell is usually one which overexpresses a TAT polypeptide. Preferably the cell is a tumor cell, e.g., a prostate, breast, ovarian, stomach, endometrial, lung, kidney, colon, colorectal, bladder cell. Various methods are available for evaluating the cellular events associated with apoptosis. For example, phosphatidyl serine (PS) translocation can be measured by annexin binding; DNA fragmentation can be evaluated through DNA laddering; and nuclear/chromatin condensation along with DNA fragmentation can be evaluated by any increase in hypodiploid cells. Preferably, the antibody, oligopeptide or other organic molecule which induces apoptosis is one which results in about 2 to 50 fold, preferably about 5 to 50 fold, and most preferably about 10 to 50 fold, induction of annexin binding relative to untreated cell in an annexin binding assay.

Antibody "effector functions" refer to those biological activities attributable to the Fc region (a native sequence Fc region or amino acid sequence variant Fc region) of an antibody, and vary with the antibody isotype. Examples of antibody effector functions include: C1q binding and complement dependent cytotoxicity; Fc receptor binding; antibody-dependent cell-mediated cytotoxicity (ADCC); phagocytosis; down regulation of cell surface receptors (e.g., B cell receptor); and B cell activation.

"Antibody-dependent cell-mediated cytotoxicity" or "ADCC" refers to a form of cytotoxicity in which secreted Ig bound onto Fc receptors (FcRs) present on certain cytotoxic cells (e.g., Natural Killer (NK)

cells, neutrophils, and macrophages) enable these cytotoxic effector cells to bind specifically to an antigen-bearing target cell and subsequently kill the target cell with cytotoxins. The antibodies "arm" the cytotoxic cells and are absolutely required for such killing. The primary cells for mediating ADCC, NK cells, express FcγRIII only, whereas monocytes express FcγRI, FcγRII and FcγRIII. FcR expression on hematopoietic cells is summarized in Table 3 on page 464 of Ravetch and Kinet, Annu. Rev. Immunol. 9:457-92 (1991). To assess ADCC activity of a molecule of interest, an in vitro ADCC assay, such as that described in US Patent No. 5,500,362 or 5,821,337 may be performed. Useful effector cells for such assays include peripheral blood mononuclear cells (PBMC) and Natural Killer (NK) cells. Alternatively, or additionally, ADCC activity of the molecule of interest may be assessed in vivo, e.g., in a animal model such as that disclosed in Clynes et al. (USA) 95:652-656 (1998).

"Fc receptor" or "FcR" describes a receptor that binds to the Fc region of an antibody. The preferred FcR is a native sequence human FcR. Moreover, a preferred FcR is one which binds an IgG antibody (a gamma receptor) and includes receptors of the FcγRI, FcγRII and FcγRIII subclasses, including allelic variants and alternatively spliced forms of these receptors. FcγRII receptors include FcγRIIA (an "activating receptor") and FcγRIIB (an "inhibiting receptor"), which have similar amino acid sequences that differ primarily in the cytoplasmic domains thereof. Activating receptor FcγRIIA contains an immunoreceptor tyrosine-based activation motif (ITAM) in its cytoplasmic domain. Inhibiting receptor FcγRIIB contains an immunoreceptor tyrosine-based inhibition motif (ITIM) in its cytoplasmic domain. (see review M. in Daëron, Annu. Rev. Immunol. 15:203-234 (1997)). FcRs are reviewed in Ravetch and Kinet, Annu. Rev. Immunol. 9:457-492 (1991); Capel et al., Immunomethods 4:25-34 (1994); and de Haas et al., J. Lab. Clin. Med. 126:330-41 (1995). Other FcRs, including those to be identified in the future, are encompassed by the term "FcR" herein. The term also includes the neonatal receptor, FcRn, which is responsible for the transfer of maternal IgGs to the fetus (Guyer et al., J. Immunol. 117:587 (1976) and Kim et al., J. Immunol. 24:249 (1994)).

"Human effector cells" are leukocytes which express one or more FcRs and perform effector functions. Preferably, the cells express at least FcγRIII and perform ADCC effector function. Examples of human leukocytes which mediate ADCC include peripheral blood mononuclear cells (PBMC), natural killer (NK) cells, monocytes, cytotoxic T cells and neutrophils; with PBMCs and NK cells being preferred. The effector cells may be isolated from a native source, e.g., from blood.

"Complement dependent cytotoxicity" or "CDC" refers to the lysis of a target cell in the presence of complement. Activation of the classical complement pathway is initiated by the binding of the first component of the complement system (C1q) to antibodies (of the appropriate subclass) which are bound to their cognate antigen. To assess complement activation, a CDC assay, e.g., as described in Gazzano-Santoro et al., J. Immunol. Methods 202:163 (1996), may be performed.

The terms "cancer" and "cancerous" refer to or describe the physiological condition in mammals that is typically characterized by unregulated cell growth. Examples of cancer include, but are not limited to, carcinoma, lymphoma, blastoma, sarcoma, and leukemia or lymphoid malignancies. More particular

examples of such cancers include squamous cell cancer (e.g., epithelial squamous cell cancer), lung cancer including small-cell lung cancer, non-small cell lung cancer, adenocarcinoma of the lung and squamous carcinoma of the lung, cancer of the peritoneum, hepatocellular cancer, gastric or stomach cancer including gastrointestinal cancer, pancreatic cancer, glioblastoma, cervical cancer, ovarian cancer, liver cancer, bladder cancer, cancer of the urinary tract, hepatoma, breast cancer, colon cancer, rectal cancer, colorectal cancer, endometrial or uterine carcinoma, salivary gland carcinoma, kidney or renal cancer, prostate cancer, vulval cancer, thyroid cancer, hepatic carcinoma, anal carcinoma, penile carcinoma, melanoma, multiple myeloma and B-cell lymphoma, brain, as well as head and neck cancer, and associated metastases.

The terms "cell proliferative disorder" and "proliferative disorder" refer to disorders that are associated with some degree of abnormal cell proliferation. In one embodiment, the cell proliferative disorder is cancer.

"Tumor", as used herein, refers to all neoplastic cell growth and proliferation, whether malignant or benign, and all pre-cancerous and cancerous cells and tissues.

An antibody, oligopeptide or other organic molecule which "induces cell death" is one which causes a viable cell to become nonviable. The cell is one which expresses a TAT polypeptide, preferably a cell that overexpresses a TAT polypeptide as compared to a normal cell of the same tissue type. The TAT polypeptide may be a transmembrane polypeptide expressed on the surface of a cancer cell or may be a polypeptide that is produced and secreted by a cancer cell. Preferably, the cell is a cancer cell, e.g., a breast, ovarian, stomach, endometrial, salivary gland, lung, kidney, colon, colorectal, thyroid, pancreatic or bladder cell. Cell death *in vitro* may be determined in the absence of complement and immune effector cells to distinguish cell death induced by antibody-dependent cell-mediated cytotoxicity (ADCC) or complement dependent cytotoxicity (CDC). Thus, the assay for cell death may be performed using heat inactivated serum (i.e., in the absence of complement) and in the absence of immune effector cells. To determine whether the antibody, oligopeptide or other organic molecule is able to induce cell death, loss of membrane integrity as evaluated by uptake of propidium iodide (PI), trypan blue (see Moore et al. Cytotechnology 17:1-11 (1995)) or 7AAD can be assessed relative to untreated cells. Preferred cell death-inducing antibodies, oligopeptides or other organic molecules are those which induce PI uptake in the PI uptake assay in BT474 cells.

A "TAT-expressing cell" is a cell which expresses an endogenous or transfected TAT polypeptide either on the cell surface or in a secreted form. A "TAT-expressing cancer" is a cancer comprising cells that have a TAT polypeptide present on the cell surface or that produce and secrete a TAT polypeptide. A "TAT-expressing cancer" optionally produces sufficient levels of TAT polypeptide on the surface of cells thereof, such that an anti-TAT antibody, oligopeptide ot other organic molecule can bind thereto and have a therapeutic effect with respect to the cancer. In another embodiment, a "TAT-expressing cancer" optionally produces and secretes sufficient levels of TAT polypeptide, such that an anti-TAT antibody, oligopeptide ot other organic molecule antagonist can bind thereto and have a therapeutic effect with respect to the cancer. With regard to the latter, the antagonist may be an antisense oligonucleotide which reduces, inhibits or prevents production and secretion of the secreted TAT polypeptide by tumor cells. A cancer which

"overexpresses" a TAT polypeptide is one which has significantly higher levels of TAT polypeptide at the cell surface thereof, or produces and secretes, compared to a noncancerous cell of the same tissue type. Such overexpression may be caused by gene amplification or by increased transcription or translation. TAT polypeptide overexpression may be determined in a diagnostic or prognostic assay by evaluating increased levels of the TAT protein present on the surface of a cell, or secreted by the cell (e.g., via an immunohistochemistry assay using anti-TAT antibodies prepared against an isolated TAT polypeptide which may be prepared using recombinant DNA technology from an isolated nucleic acid encoding the TAT polypeptide; FACS analysis, etc.). Alternatively, or additionally, one may measure levels of TAT polypeptide-encoding nucleic acid or mRNA in the cell, e.g., via fluorescent *in situ* hybridization using a nucleic acid based probe corresponding to a TAT-encoding nucleic acid or the complement thereof; (FISH; see WO98/45479 published October, 1998), Southern blotting, Northern blotting, or polymerase chain reaction (PCR) techniques, such as real time quantitative PCR (RT-PCR). One may also study TAT polypeptide overexpression by measuring shed antigen in a biological fluid such as serum, e.g, using antibody-based assays (see also, e.g., U.S. Patent No. 4,933,294 issued June 12, 1990; WO91/05264 published April 18, 1991; U.S. Patent 5,401,638 issued March 28, 1995; and Sias et al., J. Immunol. Methods 132:73-80 (1990)). Aside from the above assays, various *in vivo* assays are available to the skilled practitioner. For example, one may expose cells within the body of the patient to an antibody which is optionally labeled with a detectable label, e.g., a radioactive isotope, and binding of the antibody to cells in the patient can be evaluated, e.g., by external scanning for radioactivity or by analyzing a biopsy taken from a patient previously exposed to the antibody.

As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (i.e., is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody, oligopeptide or other organic molecule so as to generate a "labeled" antibody, oligopeptide or other organic molecule. The label may be detectable by itself (e.g. radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

The term "cytotoxic agent" as used herein refers to a substance that inhibits or prevents the function of cells and/or causes destruction of cells. The term is intended to include radioactive isotopes (e.g., At^{211}, I^{131}, I^{125}, Y^{90}, Re^{186}, Re^{188}, Sm^{153}, Bi^{212}, P^{32} and radioactive isotopes of Lu), chemotherapeutic agents e.g.

methotrexate, adriamicin, vinca alkaloids (vincristine, vinblastine, etoposide), doxorubicin, melphalan, mitomycin C, chlorambucil, daunorubicin or other intercalating agents, enzymes and fragments thereof such as nucleolytic enzymes, antibiotics, and toxins such as small molecule toxins or enzymatically active toxins of bacterial, fungal, plant or animal origin, including fragments and/or variants thereof, and the various antitumor or anticancer agents disclosed below. Other cytotoxic agents are described below. A tumoricidal agent causes destruction of tumor cells. Cytotoxins may be covalently attached to an antibody to target the toxin to a particular cell of interest which expresses the antigen. Useful cytotoxins are their linker include, without limitation, the following:

LINKERS:

MC = maleimidocaproyl

Val Cit = valine-citrulline, dipeptide site in protease cleavable linker.

Citrulline = 2-amino-5-ureido pentanoic acid

PAB = p-aminobenzylcarbamoyl ("self immolative" portion of linker)

Me = N-methyl-valine citrulline where the linker peptide bond has been modified to prevent its cleavage by cathepsin B

MC(PEG)6-OH = maleimidocaproyl- polyethylene glycol, attached to antibody cysteines.

SPP = N-Succinimidyl 4-(2-pyridylthio) pentanoate

SMCC = N-Succinimidyl 4-(N-maleimidomethyl) cyclohexane-1 carboxylate

CYTOTOXIC DRUGS:

MMAE = mono-methyl auristatin E (MW 718)

MMAF = variant of auristatin E (MMAE) with a phenylalanine at the C-terminus of the drug (MW 731.5)

AEVB = auristatin E valeryl benzylhydrazone, acid labile linker through the C-terminus of AE (MW 732)

AFP =Auristatin F phenylene diamine; (the phenylalanine variant linked to the antibody through the C-terminus via a phenylene diamine spacer) (MW 732).

A "growth inhibitory agent" when used herein refers to a compound or composition which inhibits growth of a cell, especially a TAT-expressing cancer cell, either *in vitro* or *in vivo*. Thus, the growth inhibitory agent may be one which significantly reduces the percentage of TAT-expressing cells in S phase. Examples of growth inhibitory agents include agents that block cell cycle progression (at a place other than S phase), such as agents that induce G1 arrest and M-phase arrest. Classical M-phase blockers include the vincas (vincristine and vinblastine), taxanes, and topoisomerase II inhibitors such as doxorubicin, epirubicin, daunorubicin, etoposide, and bleomycin. Those agents that arrest G1 also spill over into S-phase arrest, for example, DNA alkylating agents such as tamoxifen, prednisone, dacarbazine, mechlorethamine, cisplatin, methotrexate, 5-fluorouracil, and ara-C. Further information can be found in The Molecular Basis of Cancer, Mendelsohn and Israel, eds., Chapter 1, entitled "Cell cycle regulation, oncogenes, and antineoplastic drugs" by Murakami et al. (WB Saunders: Philadelphia, 1995), especially p. 13. The taxanes (paclitaxel and docetaxel) are anticancer

drugs both derived from the yew tree. Docetaxel (TAXOTERE®, Rhone-Poulenc Rorer), derived from the European yew, is a semisynthetic analogue of paclitaxel (TAXOL®, Bristol-Myers Squibb). Paclitaxel and docetaxel promote the assembly of microtubules from tubulin dimers and stabilize microtubules by preventing depolymerization, which results in the inhibition of mitosis in cells.

"Doxorubicin" is an anthracycline antibiotic. The full chemical name of doxorubicin is (8S-cis)-10-[(3-amino-2,3,6-trideoxy-α-L-lyxo-hexapyranosyl)oxy]-7,8,9,10-tetrahydro-6,8,11-trihydroxy-8-(hydroxyacetyl)-1-methoxy-5,12-naphthacenedione.

The term "cytokine" is a generic term for proteins released by one cell population which act on another cell as intercellular mediators. Examples of such cytokines are lymphokines, monokines, and traditional polypeptide hormones. Included among the cytokines are growth hormone such as human growth hormone, N-methionyl human growth hormone, and bovine growth hormone; parathyroid hormone; thyroxine; insulin; proinsulin; relaxin; prorelaxin; glycoprotein hormones such as follicle stimulating hormone (FSH), thyroid stimulating hormone (TSH), and luteinizing hormone (LH); hepatic growth factor; fibroblast growth factor; prolactin; placental lactogen; tumor necrosis factor-α and -β; mullerian-inhibiting substance; mouse gonadotropin-associated peptide; inhibin; activin; vascular endothelial growth factor; integrin; thrombopoietin (TPO); nerve growth factors such as NGF-β; platelet-growth factor; transforming growth factors (TGFs) such as TGF-α and TGF-β; insulin-like growth factor-I and -II; erythropoietin (EPO); osteoinductive factors; interferons such as interferon -α, -β, and -γ; colony stimulating factors (CSFs) such as macrophage-CSF (M-CSF); granulocyte-macrophage-CSF (GM-CSF); and granulocyte-CSF (G-CSF); interleukins (ILs) such as IL-1, IL- 1a, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12; a tumor necrosis factor such as TNF-α or TNF-ß; and other polypeptide factors including LIF and kit ligand (KL). As used herein, the term cytokine includes proteins from natural sources or from recombinant cell culture and biologically active equivalents of the native sequence cytokines.

The term "package insert" is used to refer to instructions customarily included in commercial packages of therapeutic products, that contain information about the indications, usage, dosage, administration, contraindications and/or warnings concerning the use of such therapeutic products.

Table 1

```
/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
 * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define  _M      -8      /* value of a match with a stop */

int      _day[26][26] = {
/*    A B C D E F G H I J K L M N O P Q R S T U V W X Y Z */
/* A */   { 2, 0,-2, 0, 0,-4, 1,-1,-1, 0,-1,-2,-1, 0,_M, 1, 0,-2, 1, 1, 0, 0,-6, 0,-3, 0},
/* B */   { 0, 3,-4, 3, 2,-5, 0, 1,-2, 0, 0,-3,-2, 2,_M,-1, 1, 0, 0, 0, 0,-2,-5, 0,-3, 1},
/* C */   {-2,-4,15,-5,-5,-4,-3,-3,-2, 0,-5,-6,-5,-4,_M,-3,-5,-4, 0,-2, 0,-2,-8, 0, 0,-5},
/* D */   { 0, 3,-5, 4, 3,-6, 1, 1,-2, 0, 0,-4,-3, 2,_M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 2},
/* E */   { 0, 2,-5, 3, 4,-5, 0, 1,-2, 0, 0,-3,-2, 1,_M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 3},
/* F */   {-4,-5,-4,-6,-5, 9,-5,-2, 1, 0,-5, 2, 0,-4,_M,-5,-5,-4,-3,-3, 0,-1, 0, 0, 7,-5},
/* G */   { 1, 0,-3, 1, 0,-5, 5,-2,-3, 0,-2,-4,-3, 0,_M,-1,-1,-3, 1, 0, 0,-1,-7, 0,-5, 0},
/* H */   {-1, 1,-3, 1, 1,-2,-2, 6,-2, 0, 0,-2,-2, 2,_M, 0, 3, 2,-1,-1, 0,-2,-3, 0, 0, 2},
/* I */   {-1,-2,-2,-2,-2, 1,-3,-2, 5, 0,-2, 2, 2,-2,_M,-2,-2,-2,-1, 0, 0, 4,-5, 0,-1,-2},
/* J */   { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */   {-1, 0,-5, 0, 0,-5,-2, 0,-2, 0, 5,-3, 0, 1,_M,-1, 1, 3, 0, 0, 0,-2,-3, 0,-4, 0},
/* L */   {-2,-3,-6,-4,-3, 2,-4,-2, 2, 0,-3, 6, 4,-3,_M,-3,-2,-3,-3,-1, 0, 2,-2, 0,-1,-2},
/* M */   {-1,-2,-5,-3,-2, 0,-3,-2, 2, 0, 0, 4, 6,-2,_M,-2,-1, 0,-2,-1, 0, 2,-4, 0,-2,-1},
/* N */   { 0, 2,-4, 2, 1,-4, 0, 2,-2, 0, 1,-3,-2, 2,_M,-1, 1, 0, 1, 0, 0,-2,-4, 0,-2, 1},
/* O */   {_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M, 0,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M},
/* P */   { 1,-1,-3,-1,-1,-5,-1, 0,-2, 0,-1,-3,-2,-1,_M, 6, 0, 0, 1, 0, 0,-1,-6, 0,-5, 0},
/* Q */   { 0, 1,-5, 2, 2,-5,-1, 3,-2, 0, 1,-2,-1, 1,_M, 0, 4, 1,-1,-1, 0,-2,-5, 0,-4, 3},
/* R */   {-2, 0,-4,-1,-1,-4,-3, 2,-2, 0, 3,-3, 0, 0,_M, 0, 1, 6, 0,-1, 0,-2, 2, 0,-4, 0},
/* S */   { 1, 0, 0, 0, 0,-3, 1,-1,-1, 0, 0,-3,-2, 1,_M, 1,-1, 0, 2, 1, 0,-1,-2, 0,-3, 0},
/* T */   { 1, 0,-2, 0, 0,-3, 0,-1, 0, 0, 0,-1,-1, 0,_M, 0,-1,-1, 1, 3, 0, 0,-5, 0,-3, 0},
/* U */   { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */   { 0,-2,-2,-2,-2,-1,-1,-2, 4, 0,-2, 2, 2,-2,_M,-1,-2,-2,-1, 0, 0, 4,-6, 0,-2,-2},
/* W */   {-6,-5,-8,-7,-7, 0,-7,-3,-5, 0,-3,-2,-4,-4,_M,-6,-5, 2,-2,-5, 0,-6,17, 0, 0,-6},
/* X */   { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */   {-3,-3, 0,-4,-4, 7,-5, 0,-1, 0,-4,-1,-2,-2,_M,-5,-4,-4,-3,-3, 0,-2, 0, 0,10,-4},
/* Z */   { 0, 1,-5, 2, 3,-5, 0, 2,-2, 0, 0,-2,-1, 1,_M, 0, 3, 0, 0, 0, 0,-2,-6, 0,-4, 4}
 };
```

Table 1 (cont')

```
/*
 */
#include <stdio.h>
#include <ctype.h>

#define  MAXJMP      16      /* max jumps in a diag */
#define  MAXGAP      24      /* don't continue to penalize gaps larger than this */
#define  JMPS        1024    /* max jmps in an path */
#define  MX          4       /* save if there's at least MX-1 bases since last jmp */

#define  DMAT        3       /* value of matching bases */
#define  DMIS        0       /* penalty for mismatched bases */
#define  DINS0       8       /* penalty for a gap */
#define  DINS1       1       /* penalty per base */
#define  PINS0       8       /* penalty for a gap */
#define  PINS1       4       /* penalty per residue */

struct jmp {
        short           n[MAXJMP];      /* size of jmp (neg for dely) */
        unsigned short  x[MAXJMP];      /* base no. of jmp in seq x */
};                                      /* limits seq to 2^16 -1 */

struct diag {
        int             score;          /* score at last jmp */
        long            offset;         /* offset of prev block */
        short           ijmp;           /* current jmp index */
        struct jmp      jp;             /* list of jmps */
};

struct path {
        int     spc;            /* number of leading spaces */
        short   n[JMPS];/* size of jmp (gap) */
        int     x[JMPS];/* loc of jmp (last elem before gap) */
};

char            *ofile;         /* output file name */
char            *namex[2];      /* seq names: getseqs() */
char            *prog;          /* prog name for err msgs */
char            *seqx[2];       /* seqs: getseqs() */
int             dmax;           /* best diag: nw() */
```

```
int         dmax0;          /* final diag */
int         dna;            /* set if dna: main() */
int         endgaps;        /* set if penalizing end gaps */
int         gapx, gapy;     /* total gaps in seqs */
int         len0, len1;     /* seq lens */
int         ngapx, ngapy;   /* total size of gaps */
int         smax;           /* max score: nw() */
int         *xbm;           /* bitmap for matching */
long        offset;         /* current offset in jmp file */
struct diag *dx;            /* holds diagonals */
struct path pp[2];          /* holds path for seqs */

char        *calloc(), *malloc(), *index(), *strcpy();
char        *getseq(), *g_calloc();
```

Table 1 (cont')

```
/* Needleman-Wunsch alignment program
 *
 * usage: progs file1 file2
 *  where file1 and file2 are two dna or two protein sequences.
 *  The sequences can be in upper- or lower-case an may contain ambiguity
 *  Any lines beginning with ';', '>' or '<' are ignored
 *  Max file length is 65535 (limited by unsigned short x in the jmp struct)
 *  A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
 *  Output is in the file "align.out"
 *
 * The program may create a tmp file in /tmp to hold info about traceback.
 * Original version developed under BSD 4.3 on a vax 8650
 */
#include "nw.h"
#include "day.h"

static  _dbval[26] = {
        1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static  _pbval[26] = {
        1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
        128, 256, 0xFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
        1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
        1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)                                                    main
        int     ac;
        char    *av[];
{
        prog = av[0];
        if (ac != 3) {
                fprintf(stderr,"usage: %s file1 file2\n", prog);
                fprintf(stderr,"where file1 and file2 are two dna or two protein sequences.\n");
                fprintf(stderr,"The sequences can be in upper- or lower-case\n");
                fprintf(stderr,"Any lines beginning with ';' or '<' are ignored\n");
                fprintf(stderr,"Output is in the file \"align.out\"\n");
                exit(1);
        }
```

```
        namex[0] = av[1];
        namex[1] = av[2];
        seqx[0] = getseq(namex[0], &len0);
        seqx[1] = getseq(namex[1], &len1);
        xbm = (dna)? _dbval : _pbval;

        endgaps = 0;                    /* 1 to penalize endgaps */
        ofile = "align.out";            /* output file */

        nw();           /* fill in the matrix, get the possible jmps */
        readjmps();     /* get the actual jmps */
        print();        /* print stats, alignment */

        cleanup(0);     /* unlink any tmp files */
}
```

Table 1 (cont')

```
/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw()                                                                    nw
{
        char        *px, *py;       /* seqs and ptrs */
        int         *ndely, *dely;  /* keep track of dely */
        int         ndelx, delx;    /* keep track of delx */
        int         *tmp;           /* for swapping row0, row1 */
        int         mis;            /* score for each type */
        int         ins0, ins1;     /* insertion penalties */
        register    id;             /* diagonal index */
        register    ij;             /* jmp index */
        register    *col0, *col1;   /* score for curr, last row */
        register    xx, yy;         /* index into seqs */

        dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

        ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
        dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
        col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
        col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
        ins0 = (dna)? DINS0 : PINS0;
        ins1 = (dna)? DINS1 : PINS1;

        smax = -10000;
        if (endgaps) {
                for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
                        col0[yy] = dely[yy] = col0[yy-1] - ins1;
                        ndely[yy] = yy;
                }
                col0[0] = 0;            /* Waterman Bull Math Biol 84 */
        }
        else
                for (yy = 1; yy <= len1; yy++)
                        dely[yy] = -ins0;
```

```
/* fill in match matrix
 */
for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
         */
        if (endgaps) {
                if (xx == 1)
                        col1[0] = delx = -(ins0+ins1);
                else
                        col1[0] = delx = col0[0] - ins1;
                ndelx = xx;
        }
        else {
                col1[0] = 0;
                delx = -ins0;
                ndelx = 0;
        }
```

Table 1 (cont')

...nw

```
for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
        mis = col0[yy-1];
        if (dna)
                mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
        else
                mis += _day[*px-'A'][*py-'A'];

        /* update penalty for del in x seq;
         * favor new del over ongong del
         * ignore MAXGAP if weighting endgaps
         */
        if (endgaps || ndely[yy] < MAXGAP) {
                if (col0[yy] - ins0 >= dely[yy]) {
                        dely[yy] = col0[yy] - (ins0+ins1);
                        ndely[yy] = 1;
                } else {
                        dely[yy] -= ins1;
                        ndely[yy]++;
                }
        } else {
                if (col0[yy] - (ins0+ins1) >= dely[yy]) {
                        dely[yy] = col0[yy] - (ins0+ins1);
                        ndely[yy] = 1;
                } else
                        ndely[yy]++;
        }

        /* update penalty for del in y seq;
         * favor new del over ongong del
         */
        if (endgaps || ndelx < MAXGAP) {
                if (col1[yy-1] - ins0 >= delx) {
                        delx = col1[yy-1] - (ins0+ins1);
                        ndelx = 1;
                } else {
                        delx -= ins1;
                        ndelx++;
                }
        } else {
```

```
            if (col1[yy-1] - (ins0+ins1) >= delx) {
                delx = col1[yy-1] - (ins0+ins1);
                ndelx = 1;
            } else
                ndelx++;
        }

        /* pick the maximum score; we're favoring
         * mis over any del and delx over dely
         */
```

Table 1 (cont')

...nw

```
                id = xx - yy + len1 - 1;
                if (mis >= delx && mis >= dely[yy])
                        col1[yy] = mis;
                else if (delx >= dely[yy]) {
                        col1[yy] = delx;
                        ij = dx[id].ijmp;
                        if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
                        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
                                dx[id].ijmp++;
                                if (++ij >= MAXJMP) {
                                        writejmps(id);
                                        ij = dx[id].ijmp = 0;
                                        dx[id].offset = offset;
                                        offset += sizeof(struct jmp) + sizeof(offset);
                                }
                        }
                        dx[id].jp.n[ij] = ndelx;
                        dx[id].jp.x[ij] = xx;
                        dx[id].score = delx;
                }
                else {
                        col1[yy] = dely[yy];
                        ij = dx[id].ijmp;
        if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
                        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
                                dx[id].ijmp++;
                                if (++ij >= MAXJMP) {
                                        writejmps(id);
                                        ij = dx[id].ijmp = 0;
                                        dx[id].offset = offset;
                                        offset += sizeof(struct jmp) + sizeof(offset);
                                }
                        }
                        dx[id].jp.n[ij] = -ndely[yy];
                        dx[id].jp.x[ij] = xx;
                        dx[id].score = dely[yy];
                }
                if (xx == len0 && yy < len1) {
                        /* last col
```

```
                                */
                                if (endgaps)
                                        col1[yy] -= ins0+ins1*(len1-yy);
                                if (col1[yy] > smax) {
                                        smax = col1[yy];
                                        dmax = id;
                                }
                        }
                }
                if (endgaps && xx < len0)
                        col1[yy-1] -= ins0+ins1*(len0-xx);
                if (col1[yy-1] > smax) {
                        smax = col1[yy-1];
                        dmax = id;
                }
                tmp = col0; col0 = col1; col1 = tmp;
        }
        (void) free((char *)ndely);
        (void) free((char *)dely);
        (void) free((char *)col0);
        (void) free((char *)col1);                      }
```

Table 1 (cont')

```
/*
 *
 * print() -- only routine visible outside this module
 *
 * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() - -put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

#include "nw.h"

#define SPC     3
#define P_LINE  256     /* maximum output line */
#define P_SPC   3       /* space between name or num and seq */

extern  _day[26][26];
int     olen;           /* set output line length */
FILE    *fx;            /* output file */

print()                                                          print
{
        int     lx, ly, firstgap, lastgap;      /* overlap */

        if ((fx = fopen(ofile, "w")) == 0) {
                fprintf(stderr,"%s: can't write %s\n", prog, ofile);
                cleanup(1);
        }
        fprintf(fx, "<first sequence: %s (length = %d)\n", namex[0], len0);
        fprintf(fx, "<second sequence: %s (length = %d)\n", namex[1], len1);
        olen = 60;
        lx = len0;
        ly = len1;
        firstgap = lastgap = 0;
        if (dmax < len1 - 1) {          /* leading gap in x */
                pp[0].spc = firstgap = len1 - dmax - 1;
```

```
            ly -= pp[0].spc;
        }
        else if (dmax > len1 - 1) {     /* leading gap in y */
            pp[1].spc = firstgap = dmax - (len1 - 1);
            lx -= pp[1].spc;
        }
        if (dmax0 < len0 - 1) {         /* trailing gap in x */
            lastgap = len0 - dmax0 -1;
            lx -= lastgap;
        }
        else if (dmax0 > len0 - 1) {    /* trailing gap in y */
            lastgap = dmax0 - (len0 - 1);
            ly -= lastgap;
        }
        getmat(lx, ly, firstgap, lastgap);
        pr_align();
    }
```

Table 1 (cont')

```
/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)                                    getmat
        int     lx, ly;                 /* "core" (minus endgaps) */
        int     firstgap, lastgap;      /* leading trailing overlap */
{
        int             nm, i0, i1, siz0, siz1;
        char            outx[32];
        double          pct;
        register        n0, n1;
        register char   *p0, *p1;

        /* get total matches, score
         */
        i0 = i1 = siz0 = siz1 = 0;
        p0 = seqx[0] + pp[1].spc;
        p1 = seqx[1] + pp[0].spc;
        n0 = pp[1].spc + 1;
        n1 = pp[0].spc + 1;

        nm = 0;
        while ( *p0 && *p1 ) {
                if (siz0) {
                        p1++;
                        n1++;
                        siz0--;
                }
                else if (siz1) {
                        p0++;
                        n0++;
                        siz1--;
                }
                else {
                        if (xbm[*p0-'A']&xbm[*p1-'A'])
                                nm++;
                        if (n0++ == pp[0].x[i0])
                                siz0 = pp[0].n[i0++];
                        if (n1++ == pp[1].x[i1])
```

```
                    siz1 = pp[1].n[i1++];
                p0++;
                p1++;
            }
        }

/* pct homology:
 * if penalizing endgaps, base is the shorter seq
 * else, knock off overhangs and take shorter core
 */
if (endgaps)
        lx = (len0 < len1)? len0 : len1;
else
        lx = (lx < ly)? lx : ly;
pct = 100.*(double)nm/(double)lx;
fprintf(fx, "\n");
fprintf(fx, "<%d match%s in an overlap of %d: %.2f percent similarity\n",
        nm, (nm == 1)? "" : "es", lx, pct);
```

Table 1 (cont')

...getmat

```
        fprintf(fx, "<gaps in first sequence: %d", gapx);
        if (gapx) {
                (void) sprintf(outx, " (%d %s%s)",
                        ngapx, (dna)? "base":"residue", (ngapx == 1)? "":"s");
                fprintf(fx,"%s", outx);

        fprintf(fx, ", gaps in second sequence: %d", gapy);
        if (gapy) {
                (void) sprintf(outx, " (%d %s%s)",
                        ngapy, (dna)? "base":"residue", (ngapy == 1)? "":"s");
                fprintf(fx,"%s", outx);
        }
        if (dna)
                fprintf(fx,
                "\n<score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
                smax, DMAT, DMIS, DINS0, DINS1);
        else
                fprintf(fx,
                "\n<score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
                smax, PINS0, PINS1);
        if (endgaps)
                fprintf(fx,
                "<endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
                firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
                lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
        else
                fprintf(fx, "<endgaps not penalized\n");
}

static          nm;             /* matches in core -- for checking */
static          lmax;           /* lengths of stripped file names */
static          ij[2];          /* jmp index for a path */
static          nc[2];          /* number at start of current line */
static          ni[2];          /* current elem number -- for gapping */
static          siz[2];
static char     *ps[2];         /* ptr to current element */
static char     *po[2];         /* ptr to next output char slot */
static char     out[2][P_LINE]; /* output line */
static char     star[P_LINE];   /* set by stars() */
```

pr_align

```
/*
 * print alignment of described in struct path pp[]
 */
static
pr_align()
{
        int             nn;     /* char count */
        int             more;
        register        i;

        for (i =  0, lmax = 0; i < 2; i++) {
                nn = stripname(namex[i]);
                if (nn > lmax)
                        lmax = nn;

                nc[i] = 1;
                ni[i] = 1;
                siz[i] = ij[i] = 0;
                ps[i] = seqx[i];
                po[i] = out[i];
}
```

Table 1 (cont')

...pr_align

```
for (nn = nm = 0, more = 1; more; ) {
        for (i = more = 0; i < 2; i++) {
                /*
                 * do we have more of this sequence?
                 */
                if (!*ps[i])
                        continue;

                more++;

                if (pp[i].spc) {        /* leading space */
                        *po[i]++ = ' ';
                        pp[i].spc--;
                }
                else if (siz[i]) {      /* in a gap */
                        *po[i]++ = '-';
                        siz[i]--;
                }
                else {                  /* we're putting a seq element
                                         */
                        *po[i] = *ps[i];
                        if (islower(*ps[i]))
                                *ps[i] = toupper(*ps[i]);
                        po[i]++;
                        ps[i]++;

                        /*
                         * are we at next gap for this seq?
                         */
                        if (ni[i] == pp[i].x[ij[i]]) {
                                /*
                                 * we need to merge all gaps
                                 * at this location
                                 */
                                siz[i] = pp[i].n[ij[i]++];
                                while (ni[i] == pp[i].x[ij[i]])
                                        siz[i] += pp[i].n[ij[i]++];
                        }
                        ni[i]++;
```

```
                        }
                }
                if (++nn == olen || !more && nn) {
                        dumpblock();
                        for (i = 0; i < 2; i++)
                                po[i] = out[i];
                        nn = 0;
                }
        }
}

/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
static
dumpblock()
{
        register  i;

        for (i = 0; i < 2; i++)
                *po[i]-- = '\0';
```

dumpblock

Table 1 (cont')

...dumpblock

```
        (void) putc('\n', fx);
        for (i = 0; i < 2; i++) {
                if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
                        if (i == 0)
                                nums(i);
                        if (i == 0 && *out[1])
                                stars();
                        putline(i);
                        if (i == 0 && *out[1])
                                fprintf(fx, star);
                        if (i == 1)
                                nums(i);
                }
        }
}

/*
 * put out a number line: dumpblock()
 */
static
```

nums

```
nums(ix)
        int     ix;     /* index in out[] holding seq line */
{
        char            nline[P_LINE];
        register        i, j;
        register char   *pn, *px, *py;

        for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
                *pn = ' ';
        for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
                if (*py == ' ' || *py == '-')
                        *pn = ' ';
                else {
                        if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                                j = (i < 0)? -i : i;
                                for (px = pn; j; j /= 10, px--)
                                        *px = j%10 + '0';
```

```
                        if (i < 0)
                                *px = '-';
                }
                else
                        *pn = ' ';
                i++;
        }
    }
    *pn = '\0';
    nc[ix] = i;
    for (pn = nline; *pn; pn++)
        (void) putc(*pn, fx);
    (void) putc('\n', fx);
}

/*
 * put out a line (name, [num], seq, [num]): dumpblock()
 */
static
putline(ix)
    int     ix;                     {
```

putline

Table 1 (cont')

...putline

```
        int             i;
        register char   *px;

        for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
                (void) putc(*px, fx);
        for (; i < lmax+P_SPC; i++)
                (void) putc(' ', fx);

        /* these count from 1:
         * ni[] is current element (from 1)
         * nc[] is number at start of current line
         */
        for (px = out[ix]; *px; px++)
                (void) putc(*px&0x7F, fx);
        (void) putc('\n', fx);
}

/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock()
 */
static
stars()
{
        int             i;
        register char   *p0, *p1, cx, *px;

        if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
            !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
                return;
        px = star;
        for (i = lmax+P_SPC; i; i--)
                *px++ = ' ';

        for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
                if (isalpha(*p0) && isalpha(*p1)) {
```

stars

```
                    if (xbm[*p0-'A']&xbm[*p1-'A']) {
                        cx = '*';
                        nm++;
                    }
                    else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
                        cx = '.';
                    else
                        cx = ' ';
                }
                else
                    cx = ' ';
                *px++ = cx;
            }
            *px++ = '\n';
            *px = '\0';
        }
```

Table 1 (cont')

stripname

```
/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
stripname(pn)
        char    *pn;    /* file name (may be path) */
{
        register char   *px, *py;

        py = 0;
        for (px = pn; *px; px++)
                if (*px == '/')
                        py = px + 1;
        if (py)
                (void) strcpy(pn, py);
        return(strlen(pn));

}
```

Table 1 (cont')

```
/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
 * readjmps() -- get the good jmps, from tmp file if necessary
 * writejmps() -- write a filled array of jmps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

char    *jname = "/tmp/homgXXXXXX";          /* tmp file for jmps */
FILE    *fj;

int     cleanup();                           /* cleanup tmp file */
long    lseek();

/*
 * remove any tmp file if we blow
 */
cleanup(i)                                                    cleanup
        int     i;
{
        if (fj)
                (void) unlink(jname);
        exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char    *
getseq(file, len)                                             getseq
        char    *file;      /* file name */
        int     *len;       /* seq len */
{
        char            line[1024], *pseq;
        register char   *px, *py;
```

```
int             natgc, tlen;
FILE            *fp;

if ((fp = fopen(file,"r")) == 0) {
        fprintf(stderr,"%s: can't read %s\n", prog, file);
        exit(1);
}
tlen = natgc = 0;
while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
                continue;
        for (px = line; *px != '\n'; px++)
                if (isupper(*px) || islower(*px))
                        tlen++;
}
if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr,"%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
}
pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
```

Table 1 (cont')

...getseq

```
        py = pseq + 4;
        *len = tlen;
        rewind(fp);

        while (fgets(line, 1024, fp)) {
                if (*line == ';' || *line == '<' || *line == '>')
                        continue;
                for (px = line; *px != '\n'; px++) {
                        if (isupper(*px))
                                *py++ = *px;
                        else if (islower(*px))
                                *py++ = toupper(*px);
                        if (index("ATGCU",*(py-1)))
                                natgc++;
                }
        }
        *py++ = '\0';
        *py = '\0';
        (void) fclose(fp);
        dna = natgc > (tlen/3);
        return(pseq+4);
}

char    *
g_calloc(msg, nx, sz)                                   g_calloc
        char    *msg;           /* program, calling routine */
        int     nx, sz;         /* number and size of elements */
{
        char            *px, *calloc();

        if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
                if (*msg) {
                        fprintf(stderr, "%s: g_calloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
                        exit(1);
                }
        }
        return(px);
}
```

readjmps

```
/*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
 */
readjmps()
{
        int             fd = -1;
        int             siz, i0, i1;
        register  i, j, xx;

        if (fj) {
                (void) fclose(fj);
                if ((fd = open(jname, O_RDONLY, 0)) < 0) {
                        fprintf(stderr, "%s: can't open() %s\n", prog, jname);
                        cleanup(1);
                }
        }
        for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
                while (1) {
                        for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                                ;
```

Table 1 (cont')

...readjmps

```
                if (j < 0 && dx[dmax].offset && fj) {
                        (void) lseek(fd, dx[dmax].offset, 0);
                        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
                        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
                        dx[dmax].ijmp = MAXJMP-1;
                }
                else
                        break;
        }
        if (i >= JMPS) {
                fprintf(stderr, "%s: too many gaps in alignment\n", prog);
                cleanup(1);
        }
        if (j >= 0) {
                siz = dx[dmax].jp.n[j];
                xx = dx[dmax].jp.x[j];
                dmax += siz;
                if (siz < 0) {                  /* gap in second seq */
                        pp[1].n[i1] = -siz;
                        xx += siz;
                        /* id = xx - yy + len1 - 1
                         */
                        pp[1].x[i1] = xx - dmax + len1 - 1;
                        gapy++;
                        ngapy -= siz;
/* ignore MAXGAP when doing endgaps */
                        siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
                        i1++;
                }
                else if (siz > 0) {     /* gap in first seq */
                        pp[0].n[i0] = siz;
                        pp[0].x[i0] = xx;
                        gapx++;
                        ngapx += siz;
/* ignore MAXGAP when doing endgaps */
                        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
                        i0++;
                }
        }
```

```
            else
                    break;
    }

    /* reverse the order of jmps
     */
    for (j = 0, i0--; j < i0; j++, i0--) {
            i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
            i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
    }
    for (j = 0, i1--; j < i1; j++, i1--) {
            i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
            i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
    }
    if (fd >= 0)
            (void) close(fd);
    if (fj) {
            (void) unlink(jname);
            fj = 0;
            offset = 0;
    }                                       }
```

Table 1 (cont')

writejmps

```
/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
writejmps(ix)
        int     ix;
{
        char    *mktemp();

        if (!fj) {
                if (mktemp(jname) < 0) {
                        fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
                        cleanup(1);
                }
                if ((fj = fopen(jname, "w")) == 0) {
                        fprintf(stderr, "%s: can't write %s\n", prog, jname);
                        exit(1);
                }
        }
        (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
        (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}
```

Table 2

TAT	XXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXYYYYYYY	(Length = 12 amino acids)

% amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the TAT polypeptide) =

5 divided by 15 = 33.3%

Table 3

TAT	XXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXYYYYYYZZYZ	(Length = 15 amino acids)

% amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the TAT polypeptide) =

5 divided by 10 = 50%

Table 4

TAT-DNA	NNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNNLLLLLLLLL	(Length = 16 nucleotides)

% nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the TAT-DNA nucleic acid sequence) =

6 divided by 14 = 42.9%

Table 5

TAT-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLVV	(Length = 9 nucleotides)

% nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the TAT-DNA nucleic acid sequence) =

4 divided by 12 = 33.3%

II. Compositions and Methods of the Invention

A. Anti-TAT Antibodies

In one embodiment, the present invention provides anti-TAT antibodies which may find use herein as therapeutic and/or diagnostic agents. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

1. Polyclonal Antibodies

Polyclonal antibodies are preferably raised in animals by multiple subcutaneous (sc) or intraperitoneal (ip) injections of the relevant antigen and an adjuvant. It may be useful to conjugate the relevant antigen (especially when synthetic peptides are used) to a protein that is immunogenic in the species to be immunized. For example, the antigen can be conjugated to keyhole limpet hemocyanin (KLH), serum albumin, bovine thyroglobulin, or soybean trypsin inhibitor, using a bifunctional or derivatizing agent, e.g., maleimidobenzoyl sulfosuccinimide ester (conjugation through cysteine residues), N-hydroxysuccinimide (through lysine residues), glutaraldehyde, succinic anhydride, $SOCl_2$, or $R^1N=C=NR$, where R and R^1 are different alkyl groups.

Animals are immunized against the antigen, immunogenic conjugates, or derivatives by combining, e.g., 100 µg or 5 µg of the protein or conjugate (for rabbits or mice, respectively) with 3 volumes of Freund's complete adjuvant and injecting the solution intradermally at multiple sites. One month later, the animals are boosted with 1/5 to 1/10 the original amount of peptide or conjugate in Freund's complete adjuvant by subcutaneous injection at multiple sites. Seven to 14 days later, the animals are bled and the serum is assayed for antibody titer. Animals are boosted until the titer plateaus. Conjugates also can be made in recombinant cell culture as protein fusions. Also, aggregating agents such as alum are suitably used to enhance the immune response.

2. Monoclonal Antibodies

Monoclonal antibodies may be made using the hybridoma method first described by Kohler et al., Nature, 256:495 (1975), or may be made by recombinant DNA methods (U.S. Patent No. 4,816,567).

In the hybridoma method, a mouse or other appropriate host animal, such as a hamster, is immunized as described above to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the protein used for immunization. Alternatively, lymphocytes may be immunized *in vitro*. After immunization, lymphocytes are isolated and then fused with a myeloma cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, pp.59-103 (Academic Press, 1986)).

The hybridoma cells thus prepared are seeded and grown in a suitable culture medium which medium preferably contains one or more substances that inhibit the growth or survival of the unfused, parental myeloma cells (also referred to as fusion partner). For example, if the parental myeloma cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the selective culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine (HAT medium), which substances prevent the growth of HGPRT-deficient cells.

Preferred fusion partner myeloma cells are those that fuse efficiently, support stable high-level production of antibody by the selected antibody-producing cells, and are sensitive to a selective medium that selects against the unfused parental cells. Preferred myeloma cell lines are murine myeloma lines, such as those derived from MOPC-21 and MPC-11 mouse tumors available from the Salk Institute Cell Distribution Center, San Diego, California USA, and SP-2 and derivatives e.g., X63-Ag8-653 cells available from the American Type Culture Collection, Manassas, Virginia, USA. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); and Brodeur et al., Monoclonal Antibody Production Techniques and Applications, pp. 51-63 (Marcel Dekker, Inc., New York, 1987)).

Culture medium in which hybridoma cells are growing is assayed for production of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunosorbent assay (ELISA).

The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis described in Munson et al., Anal. Biochem., 107:220 (1980).

Once hybridoma cells that produce antibodies of the desired specificity, affinity, and/or activity are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods (Goding, Monoclonal Antibodies: Principles and Practice, pp.59-103 (Academic Press, 1986)). Suitable culture media for this purpose include, for example, D-MEM or RPMI-1640 medium. In addition, the hybridoma cells may be grown *in vivo* as ascites tumors in an animal e.g,, by i.p. injection of the cells into mice.

The monoclonal antibodies secreted by the subclones are suitably separated from the culture medium, ascites fluid, or serum by conventional antibody purification procedures such as, for example, affinity chromatography (e.g., using protein A or protein G-Sepharose) or ion-exchange chromatography, hydroxylapatite chromatography, gel electrophoresis, dialysis, etc.

DNA encoding the monoclonal antibodies is readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as *E. coli* cells, simian COS cells, Chinese Hamster Ovary (CHO) cells, or myeloma cells that do not otherwise produce antibody protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. Review articles on recombinant expression in bacteria of DNA encoding the antibody include Skerra et al., Curr. Opinion in Immunol., 5:256-262 (1993) and Plückthun, Immunol. Revs. 130:151-188 (1992).

In a further embodiment, monoclonal antibodies or antibody fragments can be isolated from antibody phage libraries generated using the techniques described in McCafferty et al., Nature, 348:552-554 (1990). Clackson et al., Nature, 352:624-628 (1991) and Marks et al., J. Mol. Biol., 222:581-597 (1991) describe the isolation of murine and human antibodies, respectively, using phage libraries. Subsequent publications describe the production of high affinity (nM range) human antibodies by chain shuffling (Marks et al., Bio/Technology, 10:779-783 (1992)), as well as combinatorial infection and *in vivo* recombination as a strategy for constructing very large phage libraries (Waterhouse et al., Nuc. Acids. Res. 21:2265-2266 (1993)). Thus, these techniques are viable alternatives to traditional monoclonal antibody hybridoma techniques for isolation of monoclonal antibodies.

The DNA that encodes the antibody may be modified to produce chimeric or fusion antibody polypeptides, for example, by substituting human heavy chain and light chain constant domain (C_H and C_L) sequences for the homologous murine sequences (U.S. Patent No. 4,816,567; and Morrison, et al., Proc. Natl Acad. Sci. USA, 81:6851 (1984)), or by fusing the immunoglobulin coding sequence with all or part of the coding sequence for a non-immunoglobulin polypeptide (heterologous polypeptide). The non-immunoglobulin polypeptide sequences can substitute for the constant domains of an antibody, or they are substituted for the variable domains of one antigen-combining site of an antibody to create a chimeric bivalent antibody comprising one antigen-combining site having specificity for an antigen and another antigen-combining site having specificity for a different antigen.

3. Human and Humanized Antibodies

The anti-TAT antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized

antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

The choice of human variable domains, both light and heavy, to be used in making the humanized antibodies is very important to reduce antigenicity and HAMA response (human anti-mouse antibody) when the antibody is intended for human therapeutic use. According to the so-called "best-fit" method, the sequence of the variable domain of a rodent antibody is screened against the entire library of known human variable domain sequences. The human V domain sequence which is closest to that of the rodent is identified and the human framework region (FR) within it accepted for the humanized antibody (Sims et al., J. Immunol. 151:2296 (1993); Chothia et al., J. Mol. Biol., 196:901 (1987)). Another method uses a particular framework region derived from the consensus sequence of all human antibodies of a particular subgroup of light or heavy chains. The same framework may be used for several different humanized antibodies (Carter et al., Proc. Natl. Acad. Sci. USA, 89:4285 (1992); Presta et al., J. Immunol. 151:2623 (1993)).

It is further important that antibodies be humanized with retention of high binding affinity for the antigen and other favorable biological properties. To achieve this goal, according to a preferred method, humanized antibodies are prepared by a process of analysis of the parental sequences and various conceptual humanized products using three-dimensional models of the parental and humanized sequences. Three-dimensional immunoglobulin models are commonly available and are familiar to those skilled in the art. Computer programs are available which illustrate and display probable three-dimensional conformational structures of selected candidate immunoglobulin sequences. Inspection of these displays permits analysis of the likely role of the residues in the functioning of the candidate immunoglobulin sequence, i.e., the analysis of residues that influence the ability of the candidate immunoglobulin to bind its antigen. In this way, FR residues can be selected and combined from the recipient and import sequences so that the desired antibody

characteristic, such as increased affinity for the target antigen(s), is achieved. In general, the hypervariable region residues are directly and most substantially involved in influencing antigen binding.

Various forms of a humanized anti-TAT antibody are contemplated. For example, the humanized antibody may be an antibody fragment, such as a Fab, which is optionally conjugated with one or more cytotoxic agent(s) in order to generate an immunoconjugate. Alternatively, the humanized antibody may be an intact antibody, such as an intact IgG1 antibody.

As an alternative to humanization, human antibodies can be generated. For example, it is now possible to produce transgenic animals (e.g., mice) that are capable, upon immunization, of producing a full repertoire of human antibodies in the absence of endogenous immunoglobulin production. For example, it has been described that the homozygous deletion of the antibody heavy-chain joining region (J_H) gene in chimeric and germ-line mutant mice results in complete inhibition of endogenous antibody production. Transfer of the human germ-line immunoglobulin gene array into such germ-line mutant mice will result in the production of human antibodies upon antigen challenge. See, e.g., Jakobovits et al., Proc. Natl. Acad. Sci. USA, 90:2551 (1993); Jakobovits et al., Nature, 362:255-258 (1993); Bruggemann et al., Year in Immuno. 7:33 (1993); U.S. Patent Nos. 5,545,806, 5,569,825, 5,591,669 (all of GenPharm); 5,545,807; and WO 97/17852.

Alternatively, phage display technology (McCafferty et al., Nature 348:552-553 [1990]) can be used to produce human antibodies and antibody fragments *in vitro*, from immunoglobulin variable (V) domain gene repertoires from unimmunized donors. According to this technique, antibody V domain genes are cloned in-frame into either a major or minor coat protein gene of a filamentous bacteriophage, such as M13 or fd, and displayed as functional antibody fragments on the surface of the phage particle. Because the filamentous particle contains a single-stranded DNA copy of the phage genome, selections based on the functional properties of the antibody also result in selection of the gene encoding the antibody exhibiting those properties. Thus, the phage mimics some of the properties of the B-cell. Phage display can be performed in a variety of formats, reviewed in, e.g., Johnson, Kevin S. and Chiswell, David J., Current Opinion in Structural Biology 3:564-571 (1993). Several sources of V-gene segments can be used for phage display. Clackson et al., Nature, 352:624-628 (1991) isolated a diverse array of anti-oxazolone antibodies from a small random combinatorial library of V genes derived from the spleens of immunized mice. A repertoire of V genes from unimmunized human donors can be constructed and antibodies to a diverse array of antigens (including self-antigens) can be isolated essentially following the techniques described by Marks et al., J. Mol. Biol. 222:581-597 (1991), or Griffith et al., EMBO J. 12:725-734 (1993). See, also, U.S. Patent Nos. 5,565,332 and 5,573,905.

As discussed above, human antibodies may also be generated by *in vitro* activated B cells (see U.S. Patents 5,567,610 and 5,229,275).

4. Antibody fragments

In certain circumstances there are advantages of using antibody fragments, rather than whole antibodies. The smaller size of the fragments allows for rapid clearance, and may lead to improved access to solid tumors.

Various techniques have been developed for the production of antibody fragments. Traditionally, these fragments were derived via proteolytic digestion of intact antibodies (see, e.g., Morimoto et al., Journal of Biochemical and Biophysical Methods 24:107-117 (1992); and Brennan et al., Science, 229:81 (1985)). However, these fragments can now be produced directly by recombinant host cells. Fab, Fv and ScFv antibody fragments can all be expressed in and secreted from *E. coli*, thus allowing the facile production of large amounts of these fragments. Antibody fragments can be isolated from the antibody phage libraries discussed above. Alternatively, Fab'-SH fragments can be directly recovered from *E. coli* and chemically coupled to form $F(ab')_2$ fragments (Carter et al., Bio/Technology 10:163-167 (1992)). According to another approach, $F(ab')_2$ fragments can be isolated directly from recombinant host cell culture. Fab and $F(ab')_2$ fragment with increased in vivo half-life comprising a salvage receptor binding epitope residues are described in U.S. Patent No. 5,869,046. Other techniques for the production of antibody fragments will be apparent to the skilled practitioner. In other embodiments, the antibody of choice is a single chain Fv fragment (scFv). See WO 93/16185; U.S. Patent No. 5,571,894; and U.S. Patent No. 5,587,458. Fv and sFv are the only species with intact combining sites that are devoid of constant regions; thus, they are suitable for reduced nonspecific binding during in vivo use. sFv fusion proteins may be constructed to yield fusion of an effector protein at either the amino or the carboxy terminus of an sFv. See Antibody Engineering, ed. Borrebaeck, supra. The antibody fragment may also be a "linear antibody", e.g., as described in U.S. Patent 5,641,870 for example. Such linear antibody fragments may be monospecific or bispecific.

5. Bispecific Antibodies

Bispecific antibodies are antibodies that have binding specificities for at least two different epitopes. Exemplary bispecific antibodies may bind to two different epitopes of a TAT protein as described herein. Other such antibodies may combine a TAT binding site with a binding site for another protein. Alternatively, an anti-TAT arm may be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD3), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16), so as to focus and localize cellular defense mechanisms to the TAT-expressing cell. Bispecific antibodies may also be used to localize cytotoxic agents to cells which express TAT. These antibodies possess a TAT-binding arm and an arm which binds the cytotoxic agent (e.g., saporin, anti-interferon-α, vinca alkaloid, ricin A chain, methotrexate or radioactive isotope hapten). Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g., $F(ab')_2$ bispecific antibodies).

WO 96/16673 describes a bispecific anti-ErbB2/anti-FcγRIII antibody and U.S. Patent No. 5,837,234 discloses a bispecific anti-ErbB2/anti-FcγRI antibody. A bispecific anti-ErbB2/Fcα antibody is shown in WO98/02463. U.S. Patent No. 5,821,337 teaches a bispecific anti-ErbB2/anti-CD3 antibody.

Methods for making bispecific antibodies are known in the art. Traditional production of full length bispecific antibodies is based on the co-expression of two immunoglobulin heavy chain-light chain pairs, where the two chains have different specificities (Millstein et al., Nature 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of 10 different antibody molecules, of which only one has the correct bispecific structure.

Purification of the correct molecule, which is usually done by affinity chromatography steps, is rather cumbersome, and the product yields are low. Similar procedures are disclosed in WO 93/08829, and in Traunecker et al., EMBO J. 10:3655-3659 (1991).

According to a different approach, antibody variable domains with the desired binding specificities (antibody-antigen combining sites) are fused to immunoglobulin constant domain sequences. Preferably, the fusion is with an Ig heavy chain constant domain, comprising at least part of the hinge, C_H2, and C_H3 regions. It is preferred to have the first heavy-chain constant region (C_H1) containing the site necessary for light chain bonding, present in at least one of the fusions. DNAs encoding the immunoglobulin heavy chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host cell. This provides for greater flexibility in adjusting the mutual proportions of the three polypeptide fragments in embodiments when unequal ratios of the three polypeptide chains used in the construction provide the optimum yield of the desired bispecific antibody. It is, however, possible to insert the coding sequences for two or all three polypeptide chains into a single expression vector when the expression of at least two polypeptide chains in equal ratios results in high yields or when the ratios have no significant affect on the yield of the desired chain combination.

In a preferred embodiment of this approach, the bispecific antibodies are composed of a hybrid immunoglobulin heavy chain with a first binding specificity in one arm, and a hybrid immunoglobulin heavy chain-light chain pair (providing a second binding specificity) in the other arm. It was found that this asymmetric structure facilitates the separation of the desired bispecific compound from unwanted immunoglobulin chain combinations, as the presence of an immunoglobulin light chain in only one half of the bispecific molecule provides for a facile way of separation. This approach is disclosed in WO 94/04690. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology 121:210 (1986).

According to another approach described in U.S. Patent No. 5,731,168, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the C_H3 domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g., tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g., alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies include cross-linked or "heteroconjugate" antibodies. For example, one of the antibodies in the heteroconjugate can be coupled to avidin, the other to biotin. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360, WO 92/200373, and EP 03089). Heteroconjugate antibodies may be made using any convenient cross-linking methods. Suitable cross-linking agents are well known in the art, and are disclosed in U.S. Patent No. 4,676,980, along with a number of cross-linking techniques.

Techniques for generating bispecific antibodies from antibody fragments have also been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent, sodium arsenite, to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Recent progress has facilitated the direct recovery of Fab'-SH fragments from *E. coli,* which can be chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175: 217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a V_H connected to a V_L by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See Gruber et al., J. Immunol., 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., J. Immunol. 147:60 (1991).

6. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a

thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

7. Multivalent Antibodies

A multivalent antibody may be internalized (and/or catabolized) faster than a bivalent antibody by a cell expressing an antigen to which the antibodies bind. The antibodies of the present invention can be multivalent antibodies (which are other than of the IgM class) with three or more antigen binding sites (e.g. tetravalent antibodies), which can be readily produced by recombinant expression of nucleic acid encoding the polypeptide chains of the antibody. The multivalent antibody can comprise a dimerization domain and three or more antigen binding sites. The preferred dimerization domain comprises (or consists of) an Fc region or a hinge region. In this scenario, the antibody will comprise an Fc region and three or more antigen binding sites amino-terminal to the Fc region. The preferred multivalent antibody herein comprises (or consists of) three to about eight, but preferably four, antigen binding sites. The multivalent antibody comprises at least one polypeptide chain (and preferably two polypeptide chains), wherein the polypeptide chain(s) comprise two or more variable domains. For instance, the polypeptide chain(s) may comprise VD1-$(X1)_n$-VD2-$(X2)_n$-Fc, wherein VD1 is a first variable domain, VD2 is a second variable domain, Fc is one polypeptide chain of an Fc region, X1 and X2 represent an amino acid or polypeptide, and n is 0 or 1. For instance, the polypeptide chain(s) may comprise: VH-CH1-flexible linker-VH-CH1-Fc region chain; or VH-CH1-VH-CH1-Fc region chain. The multivalent antibody herein preferably further comprises at least two (and preferably four) light chain variable domain polypeptides. The multivalent antibody herein may, for instance, comprise from about two to about eight light chain variable domain polypeptides. The light chain variable domain polypeptides contemplated here comprise a light chain variable domain and, optionally, further comprise a CL domain.

8. Effector Function Engineering

It may be desirable to modify the antibody of the invention with respect to effector function, e.g., so as to enhance antigen-dependent cell-mediated cytotoxicity (ADCC) and/or complement dependent cytotoxicity (CDC) of the antibody. This may be achieved by introducing one or more amino acid substitutions in an Fc region of the antibody. Alternatively or additionally, cysteine residue(s) may be introduced in the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med. 176:1191-1195 (1992) and Shopes, B. J. Immunol. 148:2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff et al., Cancer Research 53:2560-2565 (1993). Alternatively, an antibody can be engineered which has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design 3:219-230 (1989). To increase the serum half life of the antibody, one may incorporate a salvage receptor binding epitope into the antibody (especially an antibody fragment) as described in U.S. Patent 5,739,277, for example. As used herein, the term "salvage receptor binding epitope" refers to an epitope of the Fc region of an IgG

molecule (e.g., IgG_1, IgG_2, IgG_3, or IgG_4) that is responsible for increasing the *in vivo* serum half-life of the IgG molecule.

9. Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, a growth inhibitory agent, a toxin (*e.g.*, an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re. Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

Conjugates of an antibody and one or more small molecule toxins, such as a calicheamicin, maytansinoids, a trichothene, and CC1065, and the derivatives of these toxins that have toxin activity, are also contemplated herein.

Maytansine and maytansinoids

In one preferred embodiment, an anti-TAT antibody (full length or fragments) of the invention is conjugated to one or more maytansinoid molecules.

Maytansinoids are mitototic inhibitors which act by inhibiting tubulin polymerization. Maytansine was first isolated from the east African shrub *Maytenus serrata* (U.S. Patent No. 3,896,111). Subsequently, it was discovered that certain microbes also produce maytansinoids, such as maytansinol and C-3 maytansinol esters (U.S. Patent No. 4,151,042). Synthetic maytansinol and derivatives and analogues thereof are disclosed, for example, in U.S. Patent Nos. 4,137,230; 4,248,870; 4,256,746; 4,260,608; 4,265,814; 4,294,757; 4,307,016; 4,308,268; 4,308,269; 4,309,428; 4,313,946; 4,315,929; 4,317,821; 4,322,348; 4,331,598;

4,361,650; 4,364,866; 4,424,219; 4,450,254; 4,362,663; and 4,371,533, the disclosures of which are hereby expressly incorporated by reference.

Maytansinoid-antibody conjugates

In an attempt to improve their therapeutic index, maytansine and maytansinoids have been conjugated to antibodies specifically binding to tumor cell antigens. Immunoconjugates containing maytansinoids and their therapeutic use are disclosed, for example, in U.S. Patent Nos. 5,208,020, 5,416,064 and European Patent EP 0 425 235 B1, the disclosures of which are hereby expressly incorporated by reference. Liu et al., Proc. Natl. Acad. Sci. USA 93:8618-8623 (1996) described immunoconjugates comprising a maytansinoid designated DM1 linked to the monoclonal antibody C242 directed against human colorectal cancer. The conjugate was found to be highly cytotoxic towards cultured colon cancer cells, and showed antitumor activity in an *in vivo* tumor growth assay. Chari et al., Cancer Research 52:127-131 (1992) describe immunoconjugates in which a maytansinoid was conjugated via a disulfide linker to the murine antibody A7 binding to an antigen on human colon cancer cell lines, or to another murine monoclonal antibody TA.1 that binds the HER-2/*neu* oncogene. The cytotoxicity of the TA.1-maytansonoid conjugate was tested *in vitro* on the human breast cancer cell line SK-BR-3, which expresses 3 x 10^5 HER-2 surface antigens per cell. The drug conjugate achieved a degree of cytotoxicity similar to the free maytansonid drug, which could be increased by increasing the number of maytansinoid molecules per antibody molecule. The A7-maytansinoid conjugate showed low systemic cytotoxicity in mice.

Anti-TAT polypeptide antibody-maytansinoid conjugates (immunoconjugates)

Anti-TAT antibody-maytansinoid conjugates are prepared by chemically linking an anti-TAT antibody to a maytansinoid molecule without significantly diminishing the biological activity of either the antibody or the maytansinoid molecule. An average of 3-4 maytansinoid molecules conjugated per antibody molecule has shown efficacy in enhancing cytotoxicity of target cells without negatively affecting the function or solubility of the antibody, although even one molecule of toxin/antibody would be expected to enhance cytotoxicity over the use of naked antibody. Maytansinoids are well known in the art and can be synthesized by known techniques or isolated from natural sources. Suitable maytansinoids are disclosed, for example, in U.S. Patent No. 5,208,020 and in the other patents and nonpatent publications referred to hereinabove. Preferred maytansinoids are maytansinol and maytansinol analogues modified in the aromatic ring or at other positions of the maytansinol molecule, such as various maytansinol esters.

There are many linking groups known in the art for making antibody-maytansinoid conjugates, including, for example, those disclosed in U.S. Patent No. 5,208,020 or EP Patent 0 425 235 B1, and Chari et al., Cancer Research 52:127-131 (1992). The linking groups include disufide groups, thioether groups, acid labile groups, photolabile groups, peptidase labile groups, or esterase labile groups, as disclosed in the above-identified patents, disulfide and thioether groups being preferred.

Conjugates of the antibody and maytansinoid may be made using a variety of bifunctional protein coupling agents such as N-succinimidyl-3-(2-pyridyldithio) propionate (SPDP), succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate, iminothiolane (IT), bifunctional derivatives of imidoesters (such

as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as toluene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). Particularly preferred coupling agents include N-succinimidyl-3-(2-pyridyldithio) propionate (SPDP) (Carlsson et al., Biochem. J. 173:723-737 [1978]) and N-succinimidyl-4-(2-pyridylthio)pentanoate (SPP) to provide for a disulfide linkage.

The linker may be attached to the maytansinoid molecule at various positions, depending on the type of the link. For example, an ester linkage may be formed by reaction with a hydroxyl group using conventional coupling techniques. The reaction may occur at the C-3 position having a hydroxyl group, the C-14 position modified with hyrdoxymethyl, the C-15 position modified with a hydroxyl group, and the C-20 position having a hydroxyl group. In a preferred embodiment, the linkage is formed at the C-3 position of maytansinol or a maytansinol analogue.

Calicheamicin

Another immunoconjugate of interest comprises an anti-TAT antibody conjugated to one or more calicheamicin molecules. The calicheamicin family of antibiotics are capable of producing double-stranded DNA breaks at sub-picomolar concentrations. For the preparation of conjugates of the calicheamicin family, see U.S. patents 5,712,374, 5,714,586, 5,739,116, 5,767,285, 5,770,701, 5,770,710, 5,773,001, 5,877,296 (all to American Cyanamid Company). Structural analogues of calicheamicin which may be used include, but are not limited to, γ_1^I, α_2^I, α_3^I, N-acetyl-γ_1^I, PSAG and θ^I_1 (Hinman et al., Cancer Research 53:3336-3342 (1993), Lode et al., Cancer Research 58:2925-2928 (1998) and the aforementioned U.S. patents to American Cyanamid). Another anti-tumor drug that the antibody can be conjugated is QFA which is an antifolate. Both calicheamicin and QFA have intracellular sites of action and do not readily cross the plasma membrane. Therefore, cellular uptake of these agents through antibody mediated internalization greatly enhances their cytotoxic effects.

Other cytotoxic agents

Other antitumor agents that can be conjugated to the anti-TAT antibodies of the invention include BCNU, streptozoicin, vincristine and 5-fluorouracil, the family of agents known collectively LL-E33288 complex described in U.S. patents 5,053,394, 5,770,710, as well as esperamicins (U.S. patent 5,877,296).

Enzymatically active toxins and fragments thereof which can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin and the tricothecenes. See, for example, WO 93/21232 published October 28, 1993.

The present invention further contemplates an immunoconjugate formed between an antibody and a compound with nucleolytic activity (e.g., a ribonuclease or a DNA endonuclease such as a deoxyribonuclease; DNase).

For selective destruction of the tumor, the antibody may comprise a highly radioactive atom. A variety of radioactive isotopes are available for the production of radioconjugated anti-TAT antibodies. Examples include At^{211}, I^{131}, I^{125}, Y^{90}, Re^{186}, Re^{188}, Sm^{153}, Bi^{212}, P^{32}, Pb^{212} and radioactive isotopes of Lu. When the conjugate is used for diagnosis, it may comprise a radioactive atom for scintigraphic studies, for example tc^{99m} or I^{123}, or a spin label for nuclear magnetic resonance (NMR) imaging (also known as magnetic resonance imaging, mri), such as iodine-123 again, iodine-131, indium-111, fluorine-19, carbon-13, nitrogen-15, oxygen-17, gadolinium, manganese or iron.

The radio- or other labels may be incorporated in the conjugate in known ways. For example, the peptide may be biosynthesized or may be synthesized by chemical amino acid synthesis using suitable amino acid precursors involving, for example, fluorine-19 in place of hydrogen. Labels such as tc^{99m} or I^{123}, .Re^{186}, Re^{188} and In^{111} can be attached via a cysteine residue in the peptide. Yttrium-90 can be attached via a lysine residue. The IODOGEN method (Fraker et al (1978) Biochem. Biophys. Res. Commun. 80: 49-57 can be used to incorporate iodine-123. "Monoclonal Antibodies in Immunoscintigraphy" (Chatal,CRC Press 1989) describes other methods in detail.

Conjugates of the antibody and cytotoxic agent may be made using a variety of bifunctional protein coupling agents such as N-succinimidyl-3-(2-pyridyldithio) propionate (SPDP), succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate, iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science 238:1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026. The linker may be a "cleavable linker" facilitating release of the cytotoxic drug in the cell. For example, an acid-labile linker, peptidase-sensitive linker, photolabile linker, dimethyl linker or disulfide-containing linker (Chari et al., Cancer Research 52:127-131 (1992); U.S. Patent No. 5,208,020) may be used.

Alternatively, a fusion protein comprising the anti-TAT antibody and cytotoxic agent may be made, e.g., by recombinant techniques or peptide synthesis. The length of DNA may comprise respective regions encoding the two portions of the conjugate either adjacent one another or separated by a region encoding a linker peptide which does not destroy the desired properties of the conjugate.

In yet another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for utilization in tumor pre-targeting wherein the antibody-receptor conjugate is administered to the patient,

followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) which is conjugated to a cytotoxic agent (e.g., a radionucleotide).

10. Immunoliposomes

The anti-TAT antibodies disclosed herein may also be formulated as immunoliposomes. A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688 (1985); Hwang et al., Proc. Natl Acad. Sci. USA 77:4030 (1980); U.S. Pat. Nos. 4,485,045 and 4,544,545; and WO97/38731 published October 23, 1997. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013,556.

Particularly useful liposomes can be generated by the reverse phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin et al., J. Biol. Chem. 257:286-288 (1982) via a disulfide interchange reaction. A chemotherapeutic agent is optionally contained within the liposome. See Gabizon et al., J. National Cancer Inst. 81(19):1484 (1989).

B. TAT Binding Oligopeptides

TAT binding oligopeptides of the present invention are oligopeptides that bind, preferably specifically, to a TAT polypeptide as described herein. TAT binding oligopeptides may be chemically synthesized using known oligopeptide synthesis methodology or may be prepared and purified using recombinant technology. TAT binding oligopeptides are usually at least about 5 amino acids in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 amino acids in length or more, wherein such oligopeptides that are capable of binding, preferably specifically, to a TAT polypeptide as described herein. TAT binding oligopeptides may be identified without undue experimentation using well known techniques. In this regard, it is noted that techniques for screening oligopeptide libraries for oligopeptides that are capable of specifically binding to a polypeptide target are well known in the art (see, e.g., U.S. Patent Nos. 5,556,762, 5,750,373, 4,708,871, 4,833,092, 5,223,409, 5,403,484, 5,571,689, 5,663,143; PCT Publication Nos. WO 84/03506 and WO84/03564; Geysen et al., Proc. Natl. Acad. Sci. U.S.A., 81:3998-4002 (1984); Geysen et al., Proc. Natl. Acad. Sci. U.S.A., 82:178-182 (1985); Geysen et al., in Synthetic Peptides as Antigens, 130-149 (1986); Geysen et al., J. Immunol. Meth., 102:259-274 (1987); Schoofs et al., J. Immunol., 140:611-616 (1988), Cwirla, S. E. et al. (1990) Proc. Natl. Acad. Sci. USA, 87:6378; Lowman, H.B. et al. (1991) Biochemistry, 30:10832; Clackson, T. et al. (1991) Nature, 352: 624; Marks, J. D. et al. (1991), J. Mol. Biol.,

222:581; Kang, A.S. et al. (1991) Proc. Natl. Acad. Sci. USA, 88:8363, and Smith, G. P. (1991) Current Opin. Biotechnol., 2:668).

In this regard, bacteriophage (phage) display is one well known technique which allows one to screen large oligopeptide libraries to identify member(s) of those libraries which are capable of specifically binding to a polypeptide target. Phage display is a technique by which variant polypeptides are displayed as fusion proteins to the coat protein on the surface of bacteriophage particles (Scott, J.K. and Smith, G. P. (1990) Science 249: 386). The utility of phage display lies in the fact that large libraries of selectively randomized protein variants (or randomly cloned cDNAs) can be rapidly and efficiently sorted for those sequences that bind to a target molecule with high affinity. Display of peptide (Cwirla, S. E. et al. (1990) Proc. Natl. Acad. Sci. USA, 87:6378) or protein (Lowman, H.B. et al. (1991) Biochemistry, 30:10832; Clackson, T. et al. (1991) Nature, 352: 624; Marks, J. D. et al. (1991), J. Mol. Biol., 222:581; Kang, A.S. et al. (1991) Proc. Natl. Acad. Sci. USA, 88:8363) libraries on phage have been used for screening millions of polypeptides or oligopeptides for ones with specific binding properties (Smith, G. P. (1991) Current Opin. Biotechnol., 2:668). Sorting phage libraries of random mutants requires a strategy for constructing and propagating a large number of variants, a procedure for affinity purification using the target receptor, and a means of evaluating the results of binding enrichments. U.S. Patent Nos. 5,223,409, 5,403,484, 5,571,689, and 5,663,143.

Although most phage display methods have used filamentous phage, lambdoid phage display systems (WO 95/34683; U.S. 5,627,024), T4 phage display systems (Ren, Z-J. et al. (1998) Gene 215:439; Zhu, Z. (1997) CAN 33:534; Jiang, J. et al. (1997) can 128:44380; Ren, Z-J. et al. (1997) CAN 127:215644; Ren, Z-J. (1996) Protein Sci. 5:1833; Efimov, V. P. et al. (1995) Virus Genes 10:173) and T7 phage display systems (Smith, G. P. and Scott, J.K. (1993) Methods in Enzymology, 217, 228–257; U.S. 5,766,905) are also known.

Many other improvements and variations of the basic phage display concept have now been developed. These improvements enhance the ability of display systems to screen peptide libraries for binding to selected target molecules and to display functional proteins with the potential of screening these proteins for desired properties. Combinatorial reaction devices for phage display reactions have been developed (WO 98/14277) and phage display libraries have been used to analyze and control bimolecular interactions (WO 98/20169; WO 98/20159) and properties of constrained helical peptides (WO 98/20036). WO 97/35196 describes a method of isolating an affinity ligand in which a phage display library is contacted with one solution in which the ligand will bind to a target molecule and a second solution in which the affinity ligand will not bind to the target molecule, to selectively isolate binding ligands. WO 97/46251 describes a method of biopanning a random phage display library with an affinity purified antibody and then isolating binding phage, followed by a micropanning process using microplate wells to isolate high affinity binding phage. The use of *Staphlylococcus aureus* protein A as an affinity tag has also been reported (Li et al. (1998) Mol Biotech., 9:187). WO 97/47314 describes the use of substrate subtraction libraries to distinguish enzyme specificities using a combinatorial library which may be a phage display library. A method for selecting enzymes suitable for use in detergents using phage display is described in WO 97/09446. Additional methods of selecting specific binding proteins are described in U.S. Patent Nos. 5,498,538, 5,432,018, and WO 98/15833.

Methods of generating peptide libraries and screening these libraries are also disclosed in U.S. Patent Nos. 5,723,286, 5,432,018, 5,580,717, 5,427,908, 5,498,530, 5,770,434, 5,734,018, 5,698,426, 5,763,192, and 5,723,323.

C. TAT Binding Organic Molecules

TAT binding organic molecules are organic molecules other than oligopeptides or antibodies as defined herein that bind, preferably specifically, to a TAT polypeptide as described herein. TAT binding organic molecules may be identified and chemically synthesized using known methodology (see, e.g., PCT Publication Nos. WO00/00823 and WO00/39585). TAT binding organic molecules are usually less than about 2000 daltons in size, alternatively less than about 1500, 750, 500, 250 or 200 daltons in size, wherein such organic molecules that are capable of binding, preferably specifically, to a TAT polypeptide as described herein may be identified without undue experimentation using well known techniques. In this regard, it is noted that techniques for screening organic molecule libraries for molecules that are capable of binding to a polypeptide target are well known in the art (see, e.g., PCT Publication Nos. WO00/00823 and WO00/39585). TAT binding organic molecules may be, for example, aldehydes, ketones, oximes, hydrazones, semicarbazones, carbazides, primary amines, secondary amines, tertiary amines, N-substituted hydrazines, hydrazides, alcohols, ethers, thiols, thioethers, disulfides, carboxylic acids, esters, amides, ureas, carbamates, carbonates, ketals, thioketals, acetals, thioacetals, aryl halides, aryl sulfonates, alkyl halides, alkyl sulfonates, aromatic compounds, heterocyclic compounds, anilines, alkenes, alkynes, diols, amino alcohols, oxazolidines, oxazolines, thiazolidines, thiazolines, enamines, sulfonamides, epoxides, aziridines, isocyanates, sulfonyl chlorides, diazo compounds, acid chlorides, or the like.

D. Screening for Anti-TAT Antibodies, TAT Binding Oligopeptides and TAT Binding Organic Molecules With the Desired Properties

Techniques for generating antibodies, oligopeptides and organic molecules that bind to TAT polypeptides have been described above. One may further select antibodies, oligopeptides or other organic molecules with certain biological characteristics, as desired.

The growth inhibitory effects of an anti-TAT antibody, oligopeptide or other organic molecule of the invention may be assessed by methods known in the art, e.g., using cells which express a TAT polypeptide either endogenously or following transfection with the TAT gene. For example, appropriate tumor cell lines and TAT-transfected cells may treated with an anti-TAT monoclonal antibody, oligopeptide or other organic molecule of the invention at various concentrations for a few days (e.g., 2-7) days and stained with crystal violet or MTT or analyzed by some other colorimetric assay. Another method of measuring proliferation would be by comparing ^{3}H-thymidine uptake by the cells treated in the presence or absence an anti-TAT antibody, TAT binding oligopeptide or TAT binding organic molecule of the invention. After treatment, the cells are harvested and the amount of radioactivity incorporated into the DNA quantitated in a scintillation counter. Appropriate positive controls include treatment of a selected cell line with a growth inhibitory antibody known to inhibit growth of that cell line. Growth inhibition of tumor cells *in vivo* can be determined in various ways known in the art. Preferably, the tumor cell is one that overexpresses a TAT polypeptide.

Preferably, the anti-TAT antibody, TAT binding oligopeptide or TAT binding organic molecule will inhibit cell proliferation of a TAT-expressing tumor cell *in vitro* or *in vivo* by about 25-100% compared to the untreated tumor cell, more preferably, by about 30-100%, and even more preferably by about 50-100% or 70-100%, in one embodiment, at an antibody concentration of about 0.5 to 30 μg/ml. Growth inhibition can be measured at an antibody concentration of about 0.5 to 30 μg/ml or about 0.5 nM to 200 nM in cell culture, where the growth inhibition is determined 1-10 days after exposure of the tumor cells to the antibody. The antibody is growth inhibitory *in vivo* if administration of the anti-TAT antibody at about 1 μg/kg to about 100 mg/kg body weight results in reduction in tumor size or reduction of tumor cell proliferation within about 5 days to 3 months from the first administration of the antibody, preferably within about 5 to 30 days.

To select for an anti-TAT antibody, TAT binding oligopeptide or TAT binding organic molecule which induces cell death, loss of membrane integrity as indicated by, e.g., propidium iodide (PI), trypan blue or 7AAD uptake may be assessed relative to control. A PI uptake assay can be performed in the absence of complement and immune effector cells. TAT polypeptide-expressing tumor cells are incubated with medium alone or medium containing the appropriate anti-TAT antibody (e.g, at about 10μg/ml), TAT binding oligopeptide or TAT binding organic molecule. The cells are incubated for a 3 day time period. Following each treatment, cells are washed and aliquoted into 35 mm strainer-capped 12 x 75 tubes (1ml per tube, 3 tubes per treatment group) for removal of cell clumps. Tubes then receive PI (10μg/ml). Samples may be analyzed using a FACSCAN® flow cytometer and FACSCONVERT® CellQuest software (Becton Dickinson). Those anti-TAT antibodies, TAT binding oligopeptides or TAT binding organic molecules that induce statistically significant levels of cell death as determined by PI uptake may be selected as cell death-inducing anti-TAT antibodies, TAT binding oligopeptides or TAT binding organic molecules.

To screen for antibodies, oligopeptides or other organic molecules which bind to an epitope on a TAT polypeptide bound by an antibody of interest, a routine cross-blocking assay such as that described in Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory, Ed Harlow and David Lane (1988), can be performed. This assay can be used to determine if a test antibody, oligopeptide or other organic molecule binds the same site or epitope as a known anti-TAT antibody. Alternatively, or additionally, epitope mapping can be performed by methods known in the art . For example, the antibody sequence can be mutagenized such as by alanine scanning, to identify contact residues. The mutant antibody is initailly tested for binding with polyclonal antibody to ensure proper folding. In a different method, peptides corresponding to different regions of a TAT polypeptide can be used in competition assays with the test antibodies or with a test antibody and an antibody with a characterized or known epitope.

E. Antibody Dependent Enzyme Mediated Prodrug Therapy (ADEPT)

The antibodies of the present invention may also be used in ADEPT by conjugating the antibody to a prodrug-activating enzyme which converts a prodrug (e.g., a peptidyl chemotherapeutic agent, see WO81/01145) to an active anti-cancer drug. See, for example, WO 88/07378 and U.S. Patent No. 4,975,278.

The enzyme component of the immunoconjugate useful for ADEPT includes any enzyme capable of acting on a prodrug in such a way so as to covert it into its more active, cytotoxic form.

Enzymes that are useful in the method of this invention include, but are not limited to, alkaline phosphatase useful for converting phosphate-containing prodrugs into free drugs; arylsulfatase useful for converting sulfate-containing prodrugs into free drugs; cytosine deaminase useful for converting non-toxic 5-fluorocytosine into the anti-cancer drug, 5-fluorouracil; proteases, such as serratia protease, thermolysin, subtilisin, carboxypeptidases and cathepsins (such as cathepsins B and L), that are useful for converting peptide-containing prodrugs into free drugs; D-alanylcarboxypeptidases, useful for converting prodrugs that contain D-amino acid substituents; carbohydrate-cleaving enzymes such as β-galactosidase and neuraminidase useful for converting glycosylated prodrugs into free drugs; β-lactamase useful for converting drugs derivatized with β-lactams into free drugs; and penicillin amidases, such as penicillin V amidase or penicillin G amidase, useful for converting drugs derivatized at their amine nitrogens with phenoxyacetyl or phenylacetyl groups, respectively, into free drugs. Alternatively, antibodies with enzymatic activity, also known in the art as "abzymes", can be used to convert the prodrugs of the invention into free active drugs (see, e.g., Massey, Nature 328:457-458 (1987)). Antibody-abzyme conjugates can be prepared as described herein for delivery of the abzyme to a tumor cell population.

The enzymes of this invention can be covalently bound to the anti-TAT antibodies by techniques well known in the art such as the use of the heterobifunctional crosslinking reagents discussed above. Alternatively, fusion proteins comprising at least the antigen binding region of an antibody of the invention linked to at least a functionally active portion of an enzyme of the invention can be constructed using recombinant DNA techniques well known in the art (see, e.g., Neuberger et al., Nature 312:604-608 (1984).

F. Full-Length TAT Polypeptides

The present invention also provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as TAT polypeptides. In particular, cDNAs (partial and full-length) encoding various TAT polypeptides have been identified and isolated, as disclosed in further detail in the Examples below.

As disclosed in the Examples below, various cDNA clones have been deposited with the ATCC. The actual nucleotide sequences of those clones can readily be determined by the skilled artisan by sequencing of the deposited clone using routine methods in the art. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the TAT polypeptides and encoding nucleic acids described herein, in some cases, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

G. Anti-TAT Antibody and TAT Polypeptide Variants

In addition to the anti-TAT antibodies and full-length native sequence TAT polypeptides described herein, it is contemplated that anti-TAT antibody and TAT polypeptide variants can be prepared. Anti-TAT antibody and TAT polypeptide variants can be prepared by introducing appropriate nucleotide changes into the encoding DNA, and/or by synthesis of the desired antibody or polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the anti-TAT antibody or TAT

polypeptide, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

Variations in the anti-TAT antibodies and TAT polypeptides described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the antibody or polypeptide that results in a change in the amino acid sequence as compared with the native sequence antibody or polypeptide. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the anti-TAT antibody or TAT polypeptide. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the anti-TAT antibody or TAT polypeptide with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

Anti-TAT antibody and TAT polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native antibody or protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the anti-TAT antibody or TAT polypeptide.

Anti-TAT antibody and TAT polypeptide fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating antibody or polypeptide fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired antibody or polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, anti-TAT antibody and TAT polypeptide fragments share at least one biological and/or immunological activity with the native anti-TAT antibody or TAT polypeptide disclosed herein.

In particular embodiments, conservative substitutions of interest are shown in Table 6 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 6, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 6

Original	Exemplary	Preferred

Residue	Substitutions	Substitutions
Ala (A)	val; leu; ile	val
Arg (R)	lys; gln; asn	lys
Asn (N)	gln; his; lys; arg	gln
Asp (D)	glu	glu
Cys (C)	ser	ser
Gln (Q)	asn	asn
Glu (E)	asp	asp
Gly (G)	pro; ala	ala
His (H)	asn; gln; lys; arg	arg
Ile (I)	leu; val; met; ala; phe; norleucine	leu
Leu (L)	norleucine; ile; val; met; ala; phe	ile
Lys (K)	arg; gln; asn	arg
Met (M)	leu; phe; ile	leu
Phe (F)	leu; val; ile; ala; tyr	leu
Pro (P)	ala	ala
Ser (S)	thr	thr
Thr (T)	ser	ser
Trp (W)	tyr; phe	tyr
Tyr (Y)	trp; phe; thr; ser	phe
Val (V)	ile; leu; met; phe; ala; norleucine	leu

Substantial modifications in function or immunological identity of the anti-TAT antibody or TAT polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

(1) hydrophobic: norleucine, met, ala, val, leu, ile;

(2) neutral hydrophilic: cys, ser, thr;

(3) acidic: asp, glu;

(4) basic: asn, gln, his, lys, arg;

(5) residues that influence chain orientation: gly, pro; and

(6) aromatic: trp, tyr, phe.

Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the anti-TAT antibody or TAT polypeptide variant DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, Science, 244:1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

Any cysteine residue not involved in maintaining the proper conformation of the anti-TAT antibody or TAT polypeptide also may be substituted, generally with serine, to improve the oxidative stability of the molecule and prevent aberrant crosslinking. Conversely, cysteine bond(s) may be added to the anti-TAT antibody or TAT polypeptide to improve its stability (particularly where the antibody is an antibody fragment such as an Fv fragment).

A particularly preferred type of substitutional variant involves substituting one or more hypervariable region residues of a parent antibody (e.g., a humanized or human antibody). Generally, the resulting variant(s) selected for further development will have improved biological properties relative to the parent antibody from which they are generated. A convenient way for generating such substitutional variants involves affinity maturation using phage display. Briefly, several hypervariable region sites (e.g., 6-7 sites) are mutated to generate all possible amino substitutions at each site. The antibody variants thus generated are displayed in a monovalent fashion from filamentous phage particles as fusions to the gene III product of M13 packaged within each particle. The phage-displayed variants are then screened for their biological activity (e.g., binding affinity) as herein disclosed. In order to identify candidate hypervariable region sites for modification, alanine scanning mutagenesis can be performed to identify hypervariable region residues contributing significantly to antigen binding. Alternatively, or additionally, it may be beneficial to analyze a crystal structure of the antigen-antibody complex to identify contact points between the antibody and human TAT polypeptide. Such contact residues and neighboring residues are candidates for substitution according to the techniques elaborated herein. Once such variants are generated, the panel of variants is subjected to screening as

described herein and antibodies with superior properties in one or more relevant assays may be selected for further development.

Nucleic acid molecules encoding amino acid sequence variants of the anti-TAT antibody are prepared by a variety of methods known in the art. These methods include, but are not limited to, isolation from a natural source (in the case of naturally occurring amino acid sequence variants) or preparation by oligonucleotide-mediated (or site-directed) mutagenesis, PCR mutagenesis, and cassette mutagenesis of an earlier prepared variant or a non-variant version of the anti-TAT antibody.

H. Modifications of Anti-TAT Antibodies and TAT Polypeptides

Covalent modifications of anti-TAT antibodies and TAT polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of an anti-TAT antibody or TAT polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of the anti-TAT antibody or TAT polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking anti-TAT antibody or TAT polypeptide to a water-insoluble support matrix or surface for use in the method for purifying anti-TAT antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimidate.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α-amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the anti-TAT antibody or TAT polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the antibody or polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence anti-TAT antibody or TAT polypeptide (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence anti-TAT antibody or TAT polypeptide. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

Glycosylation of antibodies and other polypeptides is typically either N-linked or O-linked. N-linked refers to the attachment of the carbohydrate moiety to the side chain of an asparagine residue. The tripeptide sequences asparagine-X-serine and asparagine-X-threonine, where X is any amino acid except proline, are the recognition sequences for enzymatic attachment of the carbohydrate moiety to the asparagine side chain. Thus, the presence of either of these tripeptide sequences in a polypeptide creates a potential glycosylation site. O-linked glycosylation refers to the attachment of one of the sugars N-aceylgalactosamine, galactose, or xylose to

a hydroxyamino acid, most commonly serine or threonine, although 5-hydroxyproline or 5-hydroxylysine may also be used.

Addition of glycosylation sites to the anti-TAT antibody or TAT polypeptide is conveniently accomplished by altering the amino acid sequence such that it contains one or more of the above-described tripeptide sequences (for N-linked glycosylation sites). The alteration may also be made by the addition of, or substitution by, one or more serine or threonine residues to the sequence of the original anti-TAT antibody or TAT polypeptide (for O-linked glycosylation sites). The anti-TAT antibody or TAT polypeptide amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the anti-TAT antibody or TAT polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the anti-TAT antibody or TAT polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the anti-TAT antibody or TAT polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of anti-TAT antibody or TAT polypeptide comprises linking the antibody or polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337. The antibody or polypeptide also may be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization (for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacylate) microcapsules, respectively), in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles and nanocapsules), or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, 16th edition, Oslo, A., Ed., (1980).

The anti-TAT antibody or TAT polypeptide of the present invention may also be modified in a way to form chimeric molecules comprising an anti-TAT antibody or TAT polypeptide fused to another, heterologous polypeptide or amino acid sequence.

In one embodiment, such a chimeric molecule comprises a fusion of the anti-TAT antibody or TAT polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the anti-TAT antibody or TAT polypeptide. The presence of such epitope-tagged forms of the anti-TAT antibody or TAT polypeptide can be

detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the anti-TAT antibody or TAT polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α-tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

In an alternative embodiment, the chimeric molecule may comprise a fusion of the anti-TAT antibody or TAT polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of an anti-TAT antibody or TAT polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH_2 and CH_3, or the hinge, CH_1, CH_2 and CH_3 regions of an IgG1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27, 1995.

I. Preparation of Anti-TAT Antibodies and TAT Polypeptides

The description below relates primarily to production of anti-TAT antibodies and TAT polypeptides by culturing cells transformed or transfected with a vector containing anti-TAT antibody- and TAT polypeptide-encoding nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare anti-TAT antibodies and TAT polypeptides. For instance, the appropriate amino acid sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the anti-TAT antibody or TAT polypeptide may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the desired anti-TAT antibody or TAT polypeptide.

1. Isolation of DNA Encoding Anti-TAT Antibody or TAT Polypeptide

DNA encoding anti-TAT antibody or TAT polypeptide may be obtained from a cDNA library prepared from tissue believed to possess the anti-TAT antibody or TAT polypeptide mRNA and to express it at

a detectable level. Accordingly, human anti-TAT antibody or TAT polypeptide DNA can be conveniently obtained from a cDNA library prepared from human tissue. The anti-TAT antibody- or TAT polypeptide-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

Libraries can be screened with probes (such as oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding anti-TAT antibody or TAT polypeptide is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

Techniques for screening a cDNA library are well known in the art. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ^{32}P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for anti-TAT antibody or TAT polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example, $CaCl_2$, $CaPO_4$, liposome-mediated and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is

generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescans*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA* ; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan*^{r}; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan*^{r}; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

Full length antibody, antibody fragments, and antibody fusion proteins can be produced in bacteria, in particular when glycosylation and Fc effector function are not needed, such as when the therapeutic antibody is conjugated to a cytotoxic agent (e.g., a toxin) and the immunoconjugate by itself shows effectiveness in tumor cell destruction. Full length antibodies have greater half life in circulation. Production in E. coli is faster and more cost efficient. For expression of antibody fragments and polypeptides in bacteria, see, e.g., U.S. 5,648,237 (Carter et. al.), U.S. 5,789,199 (Joly et al.), and U.S. 5,840,523 (Simmons et al.) which describes translation initiation regio (TIR) and signal sequences for optimizing expression and secretion, these patents

incorporated herein by reference. After expression, the antibody is isolated from the E. coli cell paste in a soluble fraction and can be purified through, e.g., a protein A or G column depending on the isotype. Final purification can be carried out similar to the process for purifying antibody expressed e.g,, in CHO cells.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for anti-TAT antibody- or TAT polypeptide-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer et al., Bio/Technology, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., J. Bacteriol., 154(2):737-742 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickeramii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilarum* (ATCC 36,906; Van den Berg et al., Bio/Technology, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus; yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., J. Basic Microbiol., 28:265-278 [1988]); *Candida; Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case et al., Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwanniomyces* such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, e.g., *Neurospora, Penicillium, Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance et al., Biochem. Biophys. Res. Commun., 112:284-289 [1983]; Tilburn et al., Gene, 26:205-221 [1983]; Yelton et al., Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 [1985]). Methylotropic yeasts are suitable herein and include, but are not limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula, Candida, Kloeckera, Pichia, Saccharomyces, Torulopsis,* and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, The Biochemistry of Methylotrophs, 269 (1982).

Suitable host cells for the expression of glycosylated anti-TAT antibody or TAT polypeptide are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as Drosophila S2 and Spodoptera Sf9, as well as plant cells, such as cell cultures of cotton, corn, potato, soybean, petunia, tomato, and tobacco. Numerous baculoviral strains and variants and corresponding permissive insect host cells from hosts such as *Spodoptera frugiperda* (caterpillar), *Aedes aegypti* (mosquito), *Aedes albopictus* (mosquito), *Drosophila melanogaster* (fruitfly), and *Bombyx mori* have been identified. A variety of viral strains for transfection are publicly available, e.g., the L-1 variant of *Autographa californica* NPV and the Bm-5 strain of *Bombyx mori* NPV, and such viruses may be used as the virus herein according to the present invention, particularly for transfection of *Spodoptera frugiperda* cells.

However, interest has been greatest in vertebrate cells, and propagation of vertebrate cells in culture (tissue culture) has become a routine procedure. Examples of useful mammalian host cell lines are monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen Virol. 36:59 (1977)); baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary cells/-DHFR (CHO, Urlaub et al., Proc. Natl. Acad. Sci. USA 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod. 23:243-251 (1980));

monkey kidney cells (CV1 ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51); TRI cells (Mather et al., Annals N.Y. Acad. Sci. 383:44-68 (1982)); MRC 5 cells; FS4 cells; and a human hepatoma line (Hep G2).

Host cells are transformed with the above-described expression or cloning vectors for anti-TAT antibody or TAT polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences.

3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding anti-TAT antibody or TAT polypeptide may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The TAT may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the anti-TAT antibody- or TAT polypeptide-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α-factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g.,

ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli.*

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the anti-TAT antibody- or TAT polypeptide-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp*1 gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp*1 gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the anti-TAT antibody- or TAT polypeptide-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β-lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (trp) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the tac promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding anti-TAT antibody or TAT polypeptide.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

Anti-TAT antibody or TAT polypeptide transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding the anti-TAT antibody or TAT polypeptide by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of

DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α-fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the anti-TAT antibody or TAT polypeptide coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding anti-TAT antibody or TAT polypeptide.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of anti-TAT antibody or TAT polypeptide in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620-625 (1981); Mantei et al., Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Culturing the Host Cells

The host cells used to produce the anti-TAT antibody or TAT polypeptide of this invention may be cultured in a variety of media. Commercially available media such as Ham's F10 (Sigma), Minimal Essential Medium ((MEM), (Sigma), RPMI-1640 (Sigma), and Dulbecco's Modified Eagle's Medium ((DMEM), Sigma) are suitable for culturing the host cells. In addition, any of the media described in Ham et al., Meth. Enz. 58:44 (1979), Barnes et al., Anal. Biochem.102:255 (1980), U.S. Pat. Nos. 4,767,704; 4,657,866; 4,927,762; 4,560,655; or 5,122,469; WO 90/03430; WO 87/00195; or U.S. Patent Re. 30,985 may be used as culture media for the host cells. Any of these media may be supplemented as necessary with hormones and/or other growth factors (such as insulin, transferrin, or epidermal growth factor), salts (such as sodium chloride, calcium, magnesium, and phosphate), buffers (such as HEPES), nucleotides (such as adenosine and thymidine), antibiotics (such as GENTAMYCIN™ drug), trace elements (defined as inorganic compounds usually present at final concentrations in the micromolar range), and glucose or an equivalent energy source. Any other necessary supplements may also be included at appropriate concentrations that would be known to those skilled in the art. The culture conditions, such as temperature, pH, and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

5. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA

hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence TAT polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to TAT DNA and encoding a specific antibody epitope.

6. Purification of Anti-TAT Antibody and TAT Polypeptide

Forms of anti-TAT antibody and TAT polypeptide may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of anti-TAT antibody and TAT polypeptide can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify anti-TAT antibody and TAT polypeptide from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the anti-TAT antibody and TAT polypeptide. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular anti-TAT antibody or TAT polypeptide produced.

When using recombinant techniques, the antibody can be produced intracellularly, in the periplasmic space, or directly secreted into the medium. If the antibody is produced intracellularly, as a first step, the particulate debris, either host cells or lysed fragments, are removed, for example, by centrifugation or ultrafiltration. Carter et al., Bio/Technology 10:163-167 (1992) describe a procedure for isolating antibodies which are secreted to the periplasmic space of *E. coli.* Briefly, cell paste is thawed in the presence of sodium acetate (pH 3.5), EDTA, and phenylmethylsulfonylfluoride (PMSF) over about 30 min. Cell debris can be removed by centrifugation. Where the antibody is secreted into the medium, supernatants from such expression systems are generally first concentrated using a commercially available protein concentration filter, for example, an Amicon or Millipore Pellicon ultrafiltration unit. A protease inhibitor such as PMSF may be included in any of the foregoing steps to inhibit proteolysis and antibiotics may be included to prevent the growth of adventitious contaminants.

The antibody composition prepared from the cells can be purified using, for example, hydroxylapatite chromatography, gel electrophoresis, dialysis, and affinity chromatography, with affinity chromatography being the preferred purification technique. The suitability of protein A as an affinity ligand depends on the species and isotype of any immunoglobulin Fc domain that is present in the antibody. Protein A can be used to purify antibodies that are based on human γ1, γ2 or γ4 heavy chains (Lindmark et al., J. Immunol. Meth. 62:1-13 (1983)). Protein G is recommended for all mouse isotypes and for human γ3 (Guss et al., EMBO J. 5:15671575 (1986)). The matrix to which the affinity ligand is attached is most often agarose, but other matrices are available. Mechanically stable matrices such as controlled pore glass or poly(styrenedivinyl)benzene allow for faster flow rates and shorter processing times than can be achieved with agarose. Where the antibody comprises a C_H3 domain, the Bakerbond ABX™resin (J. T. Baker, Phillipsburg, NJ) is useful for purification. Other techniques for protein purification such as fractionation on an ion-exchange column, ethanol precipitation, Reverse Phase HPLC, chromatography on silica, chromatography on heparin SEPHAROSE™ chromatography on an anion or cation exchange resin (such as a polyaspartic acid column), chromatofocusing, SDS-PAGE, and ammonium sulfate precipitation are also available depending on the antibody to be recovered.

Following any preliminary purification step(s), the mixture comprising the antibody of interest and contaminants may be subjected to low pH hydrophobic interaction chromatography using an elution buffer at a pH between about 2.5-4.5, preferably performed at low salt concentrations (e.g., from about 0-0.25M salt).

J. Pharmaceutical Formulations

Therapeutic formulations of the anti-TAT antibodies, TAT binding oligopeptides, TAT siRNA, TAT binding organic molecules and/or TAT polypeptides used in accordance with the present invention are prepared for storage by mixing the antibody, polypeptide, oligopeptide, siRNA or organic molecule having the desired degree of purity with optional pharmaceutically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients, or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as acetate, Tris, phosphate, citrate, and other organic acids; antioxidants including ascorbic acid and methionine; preservatives (such as octadecyldimethylbenzyl ammonium chloride; hexamethonium chloride; benzalkonium chloride, benzethonium chloride; phenol, butyl or benzyl alcohol; alkyl parabens such as methyl or propyl paraben; catechol; resorcinol; cyclohexanol; 3-pentanol; and m-cresol); low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, histidine, arginine, or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; tonicifiers such as trehalose and sodium chloride; sugars such as sucrose, mannitol, trehalose or sorbitol; surfactant such as polysorbate; salt-forming counter-ions such as sodium; metal complexes (e.g., Zn-protein complexes); and/or non-ionic surfactants such as TWEEN®, PLURONICS® or

polyethylene glycol (PEG). The antibody preferably comprises the antibody at a concentration of between 5-200 mg/ml, preferably between 10-100 mg/ml.

The formulations herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. For example, in addition to an anti-TAT antibody, TAT binding oligopeptide, TAT siRNA or TAT binding organic molecule, it may be desirable to include in the one formulation, an additional antibody, e.g., a second anti-TAT antibody which binds a different epitope on the TAT polypeptide, or an antibody to some other target such as a growth factor that affects the growth of the particular cancer. Alternatively, or additionally, the composition may further comprise a chemotherapeutic agent, cytotoxic agent, cytokine, growth inhibitory agent, anti-hormonal agent, and/or cardioprotectant. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The active ingredients may also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, 16th edition, Osol, A. Ed. (1980).

Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semi-permeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, e.g., films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and γ ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOT® (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

K. Diagnosis and Treatment with Anti-TAT Antibodies, TAT Binding Oligopeptides, TAT siRNA and TAT Binding Organic Molecules

To determine TAT expression in the cancer, various diagnostic assays are available. In one embodiment, TAT polypeptide overexpression may be analyzed by immunohistochemistry (IHC). Parrafin embedded tissue sections from a tumor biopsy may be subjected to the IHC assay and accorded a TAT protein staining intensity criteria as follows:

Score 0 - no staining is observed or membrane staining is observed in less than 10% of tumor cells.

Score 1+ - a faint/barely perceptible membrane staining is detected in more than 10% of the tumor cells. The cells are only stained in part of their membrane.

Score 2+ - a weak to moderate complete membrane staining is observed in more than 10% of the tumor cells.

Score 3+ - a moderate to strong complete membrane staining is observed in more than 10% of the tumor cells.

Those tumors with 0 or 1+ scores for TAT polypeptide expression may be characterized as not overexpressing TAT, whereas those tumors with 2+ or 3+ scores may be characterized as overexpressing TAT.

Alternatively, or additionally, FISH assays such as the INFORM® (sold by Ventana, Arizona) or PATHVISION® (Vysis, Illinois) may be carried out on formalin-fixed, paraffin-embedded tumor tissue to determine the extent (if any) of TAT overexpression in the tumor.

TAT overexpression or amplification may be evaluated using an *in vivo* diagnostic assay, e.g., by administering a molecule (such as an antibody, oligopeptide or organic molecule) which binds the molecule to be detected and is tagged with a detectable label (e.g., a radioactive isotope or a fluorescent label) and externally scanning the patient for localization of the label.

As described above, the anti-TAT antibodies, oligopeptides and organic molecules of the invention have various non-therapeutic applications. The anti-TAT antibodies, oligopeptides and organic molecules of the present invention can be useful for diagnosis and staging of TAT polypeptide-expressing cancers (e.g., in radioimaging). The antibodies, oligopeptides and organic molecules are also useful for purification or immunoprecipitation of TAT polypeptide from cells, for detection and quantitation of TAT polypeptide *in vitro*, e.g., in an ELISA or a Western blot, to kill and eliminate TAT-expressing cells from a population of mixed cells as a step in the purification of other cells.

Currently, depending on the stage of the cancer, cancer treatment involves one or a combination of the following therapies: surgery to remove the cancerous tissue, radiation therapy, and chemotherapy. Anti-TAT antibody, oligopeptide, siRNA or organic molecule therapy may be especially desirable in elderly patients who do not tolerate the toxicity and side effects of chemotherapy well and in metastatic disease where radiation therapy has limited usefulness. The tumor targeting anti-TAT antibodies, oligopeptides, siRNA and organic molecules of the invention are useful to alleviate TAT-expressing cancers upon initial diagnosis of the disease or during relapse. For therapeutic applications, the anti-TAT antibody, oligopeptide, siRNA or organic molecule can be used alone, or in combination therapy with, e.g., hormones, antiangiogens, or radiolabelled compounds, or with surgery, cryotherapy, and/or radiotherapy. Anti-TAT antibody, oligopeptide, siRNA or organic molecule treatment can be administered in conjunction with other forms of conventional therapy, either consecutively with, pre- or post-conventional therapy. Chemotherapeutic drugs such as TAXOTERE® (docetaxel), TAXOL® (palictaxel), estramustine and mitoxantrone are used in treating cancer, in particular, in good risk patients. In the present method of the invention for treating or alleviating cancer, the cancer patient can be administered anti-TAT antibody, oligopeptide, siRNA or organic molecule in conjuction with treatment with the one or more of the preceding chemotherapeutic agents. In particular, combination therapy with palictaxel and modified derivatives (see, e.g., EP0600517) is contemplated. The anti-TAT antibody, oligopeptide, siRNA or organic molecule will be administered with a therapeutically effective dose of the chemotherapeutic agent. In another embodiment, the anti-TAT antibody, oligopeptide, siRNA or organic molecule is administered in conjunction with chemotherapy to enhance the activity and efficacy of the

chemotherapeutic agent, e.g., paclitaxel. The Physicians' Desk Reference (PDR) discloses dosages of these agents that have been used in treatment of various cancers. The dosing regimen and dosages of these aforementioned chemotherapeutic drugs that are therapeutically effective will depend on the particular cancer being treated, the extent of the disease and other factors familiar to the physician of skill in the art and can be determined by the physician.

In one particular embodiment, a conjugate comprising an anti-TAT antibody, oligopeptide, siRNA or organic molecule conjugated with a cytotoxic agent is administered to the patient. Preferably, the immunoconjugate bound to the TAT protein is internalized by the cell, resulting in increased therapeutic efficacy of the immunoconjugate in killing the cancer cell to which it binds. In a preferred embodiment, the cytotoxic agent targets or interferes with the nucleic acid in the cancer cell. Examples of such cytotoxic agents are described above and include maytansinoids, calicheamicins, ribonucleases and DNA endonucleases.

The anti-TAT antibodies, oligopeptides, organic molecules or toxin conjugates thereof are administered to a human patient, in accord with known methods, such as intravenous administration, e.g.,, as a bolus or by continuous infusion over a period of time, by intramuscular, intraperitoneal, intracerobrospinal, subcutaneous, intra-articular, intrasynovial, intrathecal, oral, topical, or inhalation routes. Intravenous or subcutaneous administration of the antibody, oligopeptide or organic molecule is preferred.

Other therapeutic regimens may be combined with the administration of the anti-TAT antibody, oligopeptide or organic molecule. The combined administration includes co-administration, using separate formulations or a single pharmaceutical formulation, and consecutive administration in either order, wherein preferably there is a time period while both (or all) active agents simultaneously exert their biological activities. Preferably such combined therapy results in a synergistic therapeutic effect.

It may also be desirable to combine administration of the anti-TAT antibody or antibodies, oligopeptides or organic molecules, with administration of an antibody directed against another tumor antigen associated with the particular cancer.

In another embodiment, the therapeutic treatment methods of the present invention involves the combined administration of an anti-TAT antibody (or antibodies), oligopeptides or organic molecules and one or more chemotherapeutic agents or growth inhibitory agents, including co-administration of cocktails of different chemotherapeutic agents. Chemotherapeutic agents include estramustine phosphate, prednimustine, cisplatin, 5-fluorouracil, melphalan, cyclophosphamide, hydroxyurea and hydroxyureataxanes (such as paclitaxel and doxetaxel) and/or anthracycline antibiotics. Preparation and dosing schedules for such chemotherapeutic agents may be used according to manufacturers' instructions or as determined empirically by the skilled practitioner. Preparation and dosing schedules for such chemotherapy are also described in Chemotherapy Service Ed., M.C. Perry, Williams & Wilkins, Baltimore, MD (1992).

The antibody, oligopeptide or organic molecule may be combined with an anti-hormonal compound; e.g., an anti-estrogen compound such as tamoxifen; an anti-progesterone such as onapristone (see, EP 616 812); or an anti-androgen such as flutamide, in dosages known for such molecules. Where the cancer to be treated is androgen independent cancer, the patient may previously have been subjected to anti-androgen

therapy and, after the cancer becomes androgen independent, the anti-TAT antibody, oligopeptide or organic molecule (and optionally other agents as described herein) may be administered to the patient.

Sometimes, it may be beneficial to also co-administer a cardioprotectant (to prevent or reduce myocardial dysfunction associated with the therapy) or one or more cytokines to the patient. In addition to the above therapeutic regimes, the patient may be subjected to surgical removal of cancer cells and/or radiation therapy, before, simultaneously with, or post antibody, oligopeptide or organic molecule therapy. Suitable dosages for any of the above co-administered agents are those presently used and may be lowered due to the combined action (synergy) of the agent and anti-TAT antibody, oligopeptide or organic molecule.

For the prevention or treatment of disease, the dosage and mode of administration will be chosen by the physician according to known criteria. The appropriate dosage of antibody, oligopeptide or organic molecule will depend on the type of disease to be treated, as defined above, the severity and course of the disease, whether the antibody, oligopeptide or organic molecule is administered for preventive or therapeutic purposes, previous therapy, the patient's clinical history and response to the antibody, oligopeptide or organic molecule, and the discretion of the attending physician. The antibody, oligopeptide or organic molecule is suitably administered to the patient at one time or over a series of treatments. Preferably, the antibody, oligopeptide or organic molecule is administered by intravenous infusion or by subcutaneous injections. Depending on the type and severity of the disease, about 1 μg/kg to about 50 mg/kg body weight (e.g., about 0.1-15mg/kg/dose) of antibody can be an initial candidate dosage for administration to the patient, whether, for example, by one or more separate administrations, or by continuous infusion. A dosing regimen can comprise administering an initial loading dose of about 4 mg/kg, followed by a weekly maintenance dose of about 2 mg/kg of the anti-TAT antibody. However, other dosage regimens may be useful. A typical daily dosage might range from about 1 μg/kg to 100 mg/kg or more, depending on the factors mentioned above. For repeated administrations over several days or longer, depending on the condition, the treatment is sustained until a desired suppression of disease symptoms occurs. The progress of this therapy can be readily monitored by conventional methods and assays and based on criteria known to the physician or other persons of skill in the art.

Aside from administration of the antibody protein to the patient, the present application contemplates administration of the antibody by gene therapy. Such administration of nucleic acid encoding the antibody is encompassed by the expression "administering a therapeutically effective amount of an antibody". See, for example, WO96/07321 published March 14, 1996 concerning the use of gene therapy to generate intracellular antibodies.

There are two major approaches to getting the nucleic acid (optionally contained in a vector) into the patient's cells; *in vivo* and *ex vivo*. For *in vivo* delivery the nucleic acid is injected directly into the patient, usually at the site where the antibody is required. For *ex vivo* treatment, the patient's cells are removed, the nucleic acid is introduced into these isolated cells and the modified cells are administered to the patient either directly or, for example, encapsulated within porous membranes which are implanted into the patient (see, e.g., U.S. Patent Nos. 4,892,538 and 5,283,187). There are a variety of techniques available for introducing nucleic

acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells *in vitro*, or *in vivo* in the cells of the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation, microinjection, cell fusion, DEAE-dextran, the calcium phosphate precipitation method, etc. A commonly used vector for *ex vivo* delivery of the gene is a retroviral vector.

The currently preferred *in vivo* nucleic acid transfer techniques include transfection with viral vectors (such as adenovirus, Herpes simplex I virus, or adeno-associated virus) and lipid-based systems (useful lipids for lipid-mediated transfer of the gene are DOTMA, DOPE and DC-Chol, for example). For review of the currently known gene marking and gene therapy protocols see Anderson et al., Science 256:808-813 (1992). See also WO 93/25673 and the references cited therein.

The anti-TAT antibodies of the invention can be in the different forms encompassed by the definition of "antibody" herein. Thus, the antibodies include full length or intact antibody, antibody fragments, native sequence antibody or amino acid variants, humanized, chimeric or fusion antibodies, immunoconjugates, and functional fragments thereof. In fusion antibodies an antibody sequence is fused to a heterologous polypeptide sequence. The antibodies can be modified in the Fc region to provide desired effector functions. As discussed in more detail in the sections herein, with the appropriate Fc regions, the naked antibody bound on the cell surface can induce cytotoxicity, e.g., via antibody-dependent cellular cytotoxicity (ADCC) or by recruiting complement in complement dependent cytotoxicity, or some other mechanism. Alternatively, where it is desirable to eliminate or reduce effector function, so as to minimize side effects or therapeutic complications, certain other Fc regions may be used.

In one embodiment, the antibody competes for binding or bind substantially to, the same epitope as the antibodies of the invention. Antibodies having the biological characteristics of the present anti-TAT antibodies of the invention are also contemplated, specifically including the *in vivo* tumor targeting and any cell proliferation inhibition or cytotoxic characteristics.

Methods of producing the above antibodies are described in detail herein.

The present anti-TAT antibodies, oligopeptides, siRNAs and organic molecules are useful for treating a TAT-expressing cancer or alleviating one or more symptoms of the cancer in a mammal. Such a cancer includes prostate cancer, cancer of the urinary tract, lung cancer, breast cancer, colon cancer, colorectal cancer, and ovarian cancer, more specifically, prostate adenocarcinoma, renal cell carcinomas, colorectal adenocarcinomas, lung adenocarcinomas, lung squamous cell carcinomas, and pleural mesothelioma. The cancers encompass metastatic cancers of any of the preceding. The antibody, oligopeptide, siRNA or organic molecule is able to bind to at least a portion of the cancer cells that express TAT polypeptide in the mammal. In a preferred embodiment, the antibody, oligopeptide, siRNA or organic molecule is effective to destroy or kill TAT-expressing tumor cells or inhibit the growth of such tumor cells, *in vitro* or *in vivo*, upon binding to TAT polypeptide on the cell. Such an antibody includes a naked anti-TAT antibody (not conjugated to any agent). Naked antibodies that have cytotoxic or cell growth inhibition properties can be further harnessed with

a cytotoxic agent to render them even more potent in tumor cell destruction. Cytotoxic properties can be conferred to an anti-TAT antibody by, e.g., conjugating the antibody with a cytotoxic agent, to form an immunoconjugate as described herein. The cytotoxic agent or a growth inhibitory agent is preferably a small molecule. Toxins such as calicheamicin or a maytansinoid and analogs or derivatives thereof, are preferable.

The invention provides a composition comprising an anti-TAT antibody, oligopeptide, siRNA or organic molecule of the invention, and a carrier. For the purposes of treating cancer, compositions can be administered to the patient in need of such treatment, wherein the composition can comprise one or more anti-TAT antibodies present as an immunoconjugate or as the naked antibody. In a further embodiment, the compositions can comprise these antibodies, oligopeptides, siRNAs or organic molecules in combination with other therapeutic agents such as cytotoxic or growth inhibitory agents, including chemotherapeutic agents. The invention also provides formulations comprising an anti-TAT antibody, oligopeptide, siRNA or organic molecule of the invention, and a carrier. In one embodiment, the formulation is a therapeutic formulation comprising a pharmaceutically acceptable carrier.

Another aspect of the invention is isolated nucleic acids encoding the anti-TAT antibodies. Nucleic acids encoding both the H and L chains and especially the hypervariable region residues, chains which encode the native sequence antibody as well as variants, modifications and humanized versions of the antibody, are encompassed.

The invention also provides methods useful for treating a TAT polypeptide-expressing cancer or alleviating one or more symptoms of the cancer in a mammal, comprising administering a therapeutically effective amount of an anti-TAT antibody, oligopeptide, siRNA or organic molecule to the mammal. The antibody, oligopeptide or organic molecule therapeutic compositions can be administered short term (acute) or chronic, or intermittent as directed by physician. Also provided are methods of inhibiting the growth of, and killing a TAT polypeptide-expressing cell.

The invention also provides kits and articles of manufacture comprising at least one anti-TAT antibody, oligopeptide, siRNA or organic molecule. Kits containing anti-TAT antibodies, oligopeptides, siRNAs or organic molecules find use, e.g., for TAT cell killing assays, for purification or immunoprecipitation of TAT polypeptide from cells. For example, for isolation and purification of TAT, the kit can contain an anti-TAT antibody, oligopeptide, siRNA or organic molecule coupled to beads (e.g., sepharose beads). Kits can be provided which contain the antibodies, oligopeptides, siRNA or organic molecules for detection and quantitation of TAT *in vitro*, e.g., in an ELISA or a Western blot. Such antibody, oligopeptide, siRNA or organic molecule useful for detection may be provided with a label such as a fluorescent or radiolabel.

L. Articles of Manufacture and Kits

Another embodiment of the invention is an article of manufacture containing materials useful for the treatment of anti-TAT expressing cancer. The article of manufacture comprises a container and a label or package insert on or associated with the container. Suitable containers include, for example, bottles, vials, syringes, etc. The containers may be formed from a variety of materials such as glass or plastic. The container

holds a composition which is effective for treating the cancer condition and may have a sterile access port (for example the container may be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle). At least one active agent in the composition is an anti-TAT antibody, oligopeptide, siRNA or organic molecule of the invention. The label or package insert indicates that the composition is used for treating cancer. The label or package insert will further comprise instructions for administering the antibody, oligopeptide, siRNA or organic molecule composition to the cancer patient. Additionally, the article of manufacture may further comprise a second container comprising a pharmaceutically-acceptable buffer, such as bacteriostatic water for injection (BWFI), phosphate-buffered saline, Ringer's solution and dextrose solution. It may further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, needles, and syringes.

Kits are also provided that are useful for various purposes , e.g., for TAT-expressing cell killing assays, for purification or immunoprecipitation of TAT polypeptide from cells. For isolation and purification of TAT polypeptide, the kit can contain an anti-TAT antibody, oligopeptide, siRNA or organic molecule coupled to beads (e.g., sepharose beads). Kits can be provided which contain the antibodies, oligopeptides, siRNAs or organic molecules for detection and quantitation of TAT polypeptide *in vitro*, e.g., in an ELISA or a Western blot. As with the article of manufacture, the kit comprises a container and a label or package insert on or associated with the container. The container holds a composition comprising at least one anti-TAT antibody, oligopeptide, siRNA or organic molecule of the invention. Additional containers may be included that contain, e.g., diluents and buffers, control antibodies. The label or package insert may provide a description of the composition as well as instructions for the intended in vitro or diagnostic use.

M. Uses for TAT Polypeptides and TAT-Polypeptide Encoding Nucleic Acids

Nucleotide sequences (or their complement) encoding TAT polypeptides have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA, siRNA and DNA probes. TAT-encoding nucleic acid will also be useful for the preparation of TAT polypeptides by the recombinant techniques described herein, wherein those TAT polypeptides may find use, for example, in the preparation of anti-TAT antibodies as described herein.

The full-length native sequence TAT gene, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length TAT cDNA or to isolate still other cDNAs (for instance, those encoding naturally-occurring variants of TAT or TAT from other species) which have a desired sequence identity to the native TAT sequence disclosed herein. Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from at least partially novel regions of the full length native nucleotide sequence wherein those regions may be determined without undue experimentation or from genomic sequences including promoters, enhancer elements and introns of native sequence TAT. By way of example, a screening method will comprise isolating the coding region of the TAT gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S, or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of

the TAT gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below. Any EST sequences disclosed in the present application may similarly be employed as probes, using the methods disclosed herein.

Other useful fragments of the TAT-encoding nucleic acids include antisense or sense oligonucleotides comprising a singe-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target TAT mRNA (sense) or TAT DNA (antisense) sequences. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of TAT DNA. Such a fragment generally comprises at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (Cancer Res. 48:2659, 1988) and van der Krol et al. (BioTechniques 6:958, 1988).

Binding of antisense or sense oligonucleotides to target nucleic acid sequences results in the formation of duplexes that block transcription or translation of the target sequence by one of several means, including enhanced degradation of the duplexes, premature termination of transcription or translation, or by other means. Such methods are encompassed by the present invention. The antisense oligonucleotides thus may be used to block expression of TAT proteins, wherein those TAT proteins may play a role in the induction of cancer in mammals. Antisense or sense oligonucleotides further comprise oligonucleotides having modified sugar-phosphodiester backbones (or other sugar linkages, such as those described in WO 91/06629) and wherein such sugar linkages are resistant to endogenous nucleases. Such oligonucleotides with resistant sugar linkages are stable *in vivo* (i.e., capable of resisting enzymatic degradation) but retain sequence specificity to be able to bind to target nucleotide sequences.

Preferred intragenic sites for antisense binding include the region incorporating the translation initiation/start codon (5'-AUG / 5'-ATG) or termination/stop codon (5'-UAA, 5'-UAG and 5-UGA / 5'-TAA, 5'-TAG and 5'-TGA) of the open reading frame (ORF) of the gene. These regions refer to a portion of the mRNA or gene that encompasses from about 25 to about 50 contiguous nucleotides in either direction (i.e., 5' or 3') from a translation initiation or termination codon. Other preferred regions for antisense binding include: introns; exons; intron-exon junctions; the open reading frame (ORF) or "coding region," which is the region between the translation initiation codon and the translation termination codon; the 5' cap of an mRNA which comprises an N7-methylated guanosine residue joined to the 5'-most residue of the mRNA via a 5'-5' triphosphate linkage and includes 5' cap structure itself as well as the first 50 nucleotides adjacent to the cap; the 5' untranslated region (5'UTR), the portion of an mRNA in the 5' direction from the translation initiation codon, and thus including nucleotides between the 5' cap site and the translation initiation codon of an mRNA or corresponding nucleotides on the gene; and the 3' untranslated region (3'UTR), the portion of an mRNA in the 3' direction from the translation termination codon, and thus including nucleotides between the translation termination codon and 3' end of an mRNA or corresponding nucleotides on the gene.

Specific examples of preferred antisense compounds useful for inhibiting expression of TAT proteins include oligonucleotides containing modified backbones or non-natural internucleoside linkages. Oligonucleotides having modified backbones include those that retain a phosphorus atom in the backbone and those that do not have a phosphorus atom in the backbone. For the purposes of this specification, and as sometimes referenced in the art, modified oligonucleotides that do not have a phosphorus atom in their internucleoside backbone can also be considered to be oligonucleosides. Preferred modified oligonucleotide backbones include, for example, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotri-esters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates, 5'-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, selenophosphates and borano-phosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity wherein one or more internucleotide linkages is a 3' to 3', 5' to 5' or 2' to 2' linkage. Preferred oligonucleotides having inverted polarity comprise a single 3' to 3' linkage at the 3'-most internucleotide linkage i.e. a single inverted nucleoside residue which may be abasic (the nucleobase is missing or has a hydroxyl group in place thereof). Various salts, mixed salts and free acid forms are also included. Representative United States patents that teach the preparation of phosphorus-containing linkages include, but are not limited to, U.S. Pat. Nos.: 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; 5,194,599; 5,565,555; 5,527,899; 5,721,218; 5,672,697 and 5,625,050, each of which is herein incorporated by reference.

Preferred modified oligonucleotide backbones that do not include a phosphorus atom therein have backbones that are formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom and alkyl or cycloalkyl internucleoside linkages, or one or more short chain heteroatomic or heterocyclic internucleoside linkages. These include those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfoxide and sulfone backbones; formacetyl and thioformacetyl backbones; methylene formacetyl and thioformacetyl backbones; riboacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH.sub.2 component parts. Representative United States patents that teach the preparation of such oligonucleosides include, but are not limited to,. U.S. Pat. Nos.: 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437; 5,792,608; 5,646,269 and 5,677,439, each of which is herein incorporated by reference.

In other preferred antisense oligonucleotides, both the sugar and the internucleoside linkage, i.e., the backbone, of the nucleotide units are replaced with novel groups. The base units are maintained for hybridization with an appropriate nucleic acid target compound. One such oligomeric compound, an

oligonucleotide mimetic that has been shown to have excellent hybridization properties, is referred to as a peptide nucleic acid (PNA). In PNA compounds, the sugar-backbone of an oligonucleotide is replaced with an amide containing backbone, in particular an aminoethylglycine backbone. The nucleobases are retained and are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone. Representative United States patents that teach the preparation of PNA compounds include, but are not limited to, U.S. Pat. Nos.: 5,539,082; 5,714,331; and 5,719,262, each of which is herein incorporated by reference. Further teaching of PNA compounds can be found in Nielsen et al., Science, 1991, 254, 1497-1500.

Preferred antisense oligonucleotides incorporate phosphorothioate backbones and/or heteroatom backbones, and in particular -CH_2-NH-O-CH_2-, -CH_2-N(CH_3)-O-CH_2- [known as a methylene (methylimino) or MMI backbone], -CH_2-O-N(CH_3)-CH_2-, -CH_2-N(CH_3)-N(CH_3)-CH_2- and -O-N(CH_3)-CH_2-CH_2- [wherein the native phosphodiester backbone is represented as -O-P-O-CH_2-] described in the above referenced U.S. Pat. No. 5,489,677, and the amide backbones of the above referenced U.S. Pat. No. 5,602,240. Also preferred are antisense oligonucleotides having morpholino backbone structures of the above-referenced U.S. Pat. No. 5,034,506.

Modified oligonucleotides may also contain one or more substituted sugar moieties. Preferred oligonucleotides comprise one of the following at the 2' position: OH; F; O-alkyl, S-alkyl, or N-alkyl; O-alkenyl, S-alkeynyl, or N-alkenyl; O-alkynyl, S-alkynyl or N-alkynyl; or O-alkyl-O-alkyl, wherein the alkyl, alkenyl and alkynyl may be substituted or unsubstituted C_1 to C_{10} alkyl or C_2 to C_{10} alkenyl and alkynyl. Particularly preferred are $O[(CH_2)_nO]_mCH_3$, $O(CH_2)_nOCH_3$, $O(CH_2)_nNH_2$, $O(CH_2)_nCH_3$, $O(CH_2)_nONH_2$, and $O(CH_2)_nON[(CH_2)_nCH_3)]_2$, where n and m are from 1 to about 10. Other preferred antisense oligonucleotides comprise one of the following at the 2' position: C_1 to C_{10} lower alkyl, substituted lower alkyl, alkenyl, alkynyl, alkaryl, aralkyl, O-alkaryl or O-aralkyl, SH, SCH_3, OCN, Cl, Br, CN, CF_3, OCF_3, $SOCH_3$, SO_2 CH_3, ONO_2, NO_2, N_3, NH_2, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, an RNA cleaving group, a reporter group, an intercalator, a group for improving the pharmacokinetic properties of an oligonucleotide, or a group for improving the pharmacodynamic properties of an oligonucleotide, and other substituents having similar properties. A preferred modification includes 2'-methoxyethoxy (2'-O-$CH_2CH_2OCH_3$, also known as 2'-O-(2-methoxyethyl) or 2'-MOE) (Martin et al., Helv. Chim. Acta, 1995, 78, 486-504) i.e., an alkoxyalkoxy group. A further preferred modification includes 2'-dimethylaminooxyethoxy, i.e., a $O(CH_2)_2ON(CH_3)_2$ group, also known as 2'-DMAOE, as described in examples hereinbelow, and 2'-dimethylaminoethoxyethoxy (also known in the art as 2'-O-dimethylaminoethoxyethyl or 2'-DMAEOE), i.e., 2'-O-CH_2-O-CH_2-N(CH_2).

A further prefered modification includes Locked Nucleic Acids (LNAs) in which the 2'-hydroxyl group is linked to the 3' or 4' carbon atom of the sugar ring thereby forming a bicyclic sugar moiety. The linkage is preferably a methelyne (-CH_2-)$_n$ group bridging the 2' oxygen atom and the 4' carbon atom wherein n is 1 or 2. LNAs and preparation thereof are described in WO 98/39352 and WO 99/14226.

Other preferred modifications include 2'-methoxy (2'-O-CH_3), 2'-aminopropoxy (2'-$OCH_2CH_2CH_2$ NH_2), 2'-allyl (2'-CH_2-CH=CH_2), 2'-O-allyl (2'-O-CH_2-CH=CH_2) and 2'-fluoro (2'-F). The 2'-modification may

be in the arabino (up) position or ribo (down) position. A preferred 2'-arabino modification is 2'-F. Similar modifications may also be made at other positions on the oligonucleotide, particularly the 3' position of the sugar on the 3' terminal nucleotide or in 2'-5' linked oligonucleotides and the 5' position of 5' terminal nucleotide. Oligonucleotides may also have sugar mimetics such as cyclobutyl moieties in place of the pentofuranosyl sugar. Representative United States patents that teach the preparation of such modified sugar structures include, but are not limited to, U.S. Pat. Nos.: 4,981,957; 5,118,800; 5,319,080; 5,359,044; 5,393,878; 5,446,137; 5,466,786; 5,514,785; 5,519,134; 5,567,811; 5,576,427; 5,591,722; 5,597,909; 5,610,300; 5,627,053; 5,639,873; 5,646,265; 5,658,873; 5,670,633; 5,792,747; and 5,700,920, each of which is herein incorporated by reference in its entirety.

Oligonucleotides may also include nucleobase (often referred to in the art simply as "base") modifications or substitutions. As used herein, "unmodified" or "natural" nucleobases include the purine bases adenine (A) and guanine (G), and the pyrimidine bases thymine (T), cytosine (C) and uracil (U). Modified nucleobases include other synthetic and natural nucleobases such as 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-halouracil and cytosine, 5-propynyl (-C≡C-CH_3 or -CH_2-C≡CH) uracil and cytosine and other alkynyl derivatives of pyrimidine bases, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 2-F-adenine, 2-amino-adenine, 8-azaguanine and 8-azaadenine, 7-deazaguanine and 7-deazaadenine and 3-deazaguanine and 3-deazaadenine. Further modified nucleobases include tricyclic pyrimidines such as phenoxazine cytidine(1H-pyrimido[5,4-b][1,4]benzoxazin-2(3H)-one), phenothiazine cytidine (1H-pyrimido[5,4-b][1,4]benzothiazin-2(3H)-one), G-clamps such as a substituted phenoxazine cytidine (e.g. 9-(2-aminoethoxy)-H-pyrimido[5,4-b][1,4]benzoxazin-2(3H)-one), carbazole cytidine (2H-pyrimido[4,5-b]indol-2-one), pyridoindole cytidine (H-pyrido[3',2':4,5]pyrrolo[2,3-d]pyrimidin-2-one). Modified nucleobases may also include those in which the purine or pyrimidine base is replaced with other heterocycles, for example 7-deaza-adenine, 7-deazaguanosine, 2-aminopyridine and 2-pyridone. Further nucleobases include those disclosed in U.S. Pat. No. 3,687,808, those disclosed in The Concise Encyclopedia Of Polymer Science And Engineering, pages 858-859, Kroschwitz, J. I., ed. John Wiley & Sons, 1990, and those disclosed by Englisch et al., Angewandte Chemie, International Edition, 1991, 30, 613. Certain of these nucleobases are particularly useful for increasing the binding affinity of the oligomeric compounds of the invention. These include 5-substituted pyrimidines, 6-azapyrimidines and N-2, N-6 and O-6 substituted purines, including 2-aminopropyladenine, 5-propynyluracil and 5-propynylcytosine. 5-methylcytosine substitutions have been shown to increase nucleic acid duplex stability by 0.6-1.2.degree. C. (Sanghvi et al, Antisense Research and Applications, CRC Press, Boca Raton, 1993, pp. 276-278) and are preferred base substitutions, even more particularly when combined with 2'-O-methoxyethyl sugar modifications. Representative United States patents that teach the preparation of modified nucleobases include, but are not

limited to: U.S. Pat. No. 3,687,808, as well as U.S. Pat. Nos.: 4,845,205; 5,130,302; 5,134,066; 5,175,273; 5,367,066; 5,432,272; 5,457,187; 5,459,255; 5,484,908; 5,502,177; 5,525,711; 5,552,540; 5,587,469; 5,594,121, 5,596,091; 5,614,617; 5,645,985; 5,830,653; 5,763,588; 6,005,096; 5,681,941 and 5,750,692, each of which is herein incorporated by reference.

Another modification of antisense oligonucleotides chemically linking to the oligonucleotide one or more moieties or conjugates which enhance the activity, cellular distribution or cellular uptake of the oligonucleotide. The compounds of the invention can include conjugate groups covalently bound to functional groups such as primary or secondary hydroxyl groups. Conjugate groups of the invention include intercalators, reporter molecules, polyamines, polyamides, polyethylene glycols, polyethers, groups that enhance the pharmacodynamic properties of oligomers, and groups that enhance the pharmacokinetic properties of oligomers. Typical conjugates groups include cholesterols, lipids, cation lipids, phospholipids, cationic phospholipids, biotin, phenazine, folate, phenanthridine, anthraquinone, acridine, fluoresceins, rhodamines, coumarins, and dyes. Groups that enhance the pharmacodynamic properties, in the context of this invention, include groups that improve oligomer uptake, enhance oligomer resistance to degradation, and/or strengthen sequence-specific hybridization with RNA. Groups that enhance the pharmacokinetic properties, in the context of this invention, include groups that improve oligomer uptake, distribution, metabolism or excretion. Conjugate moieties include but are not limited to lipid moieties such as a cholesterol moiety (Letsinger et al., Proc. Natl. Acad. Sci. USA, 1989, 86, 6553-6556), cholic acid (Manoharan et al., Bioorg. Med. Chem. Let., 1994, 4, 1053-1060), a thioether, e.g., hexyl-S-tritylthiol (Manoharan et al., Ann. N.Y. Acad. Sci., 1992, 660, 306-309; Manoharan et al., Bioorg. Med. Chem. Let., 1993, 3, 2765-2770), a thiocholesterol (Oberhauser et al., Nucl. Acids Res., 1992, 20, 533-538), an aliphatic chain, e.g., dodecandiol or undecyl residues (Saison-Behmoaras et al., EMBO J., 1991, 10, 1111-1118; Kabanov et al., FEBS Lett., 1990, 259, 327-330; Svinarchuk et al., Biochimie, 1993, 75, 49-54), a phospholipid, e.g., di-hexadecyl-rac-glycerol or triethyl-ammonium 1,2-di-O-hexadecyl-rac-glycero-3-H-phosphonate (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651-3654; Shea et al., Nucl. Acids Res., 1990, 18, 3777-3783), a polyamine or a polyethylene glycol chain (Manoharan et al., Nucleosides & Nucleotides, 1995, 14, 969-973), or adamantane acetic acid (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651-3654), a palmityl moiety (Mishra et al., Biochim. Biophys. Acta, 1995, 1264, 229-237), or an octadecylamine or hexylamino-carbonyl-oxycholesterol moiety. Oligonucleotides of the invention may also be conjugated to active drug substances, for example, aspirin, warfarin, phenylbutazone, ibuprofen, suprofen, fenbufen, ketoprofen, (S)-(+)-pranoprofen, carprofen, dansylsarcosine, 2,3,5-triiodobenzoic acid, flufenamic acid, folinic acid, a benzothiadiazide, chlorothiazide, a diazepine, indomethicin, a barbiturate, a cephalosporin, a sulfa drug, an antidiabetic, an antibacterial or an antibiotic. Oligonucleotide-drug conjugates and their preparation are described in U.S. patent application Ser. No. 09/334,130 (filed Jun. 15, 1999) and United States patents Nos.: 4,828,979; 4,948,882; 5,218,105; 5,525,465; 5,541,313; 5,545,730; 5,552,538; 5,578,717, 5,580,731; 5,580,731; 5,591,584; 5,109,124; 5,118,802; 5,138,045; 5,414,077; 5,486,603; 5,512,439; 5,578,718; 5,608,046; 4,587,044; 4,605,735; 4,667,025; 4,762,779; 4,789,737; 4,824,941; 4,835,263; 4,876,335; 4,904,582; 4,958,013; 5,082,830;

5,112,963; 5,214,136; 5,082,830; 5,112,963; 5,214,136; 5,245,022; 5,254,469; 5,258,506; 5,262,536; 5,272,250; 5,292,873; 5,317,098; 5,371,241, 5,391,723; 5,416,203, 5,451,463; 5,510,475; 5,512,667; 5,514,785; 5,565,552; 5,567,810; 5,574,142; 5,585,481; 5,587,371; 5,595,726; 5,597,696; 5,599,923; 5,599,928 and 5,688,941, each of which is herein incorporated by reference.

It is not necessary for all positions in a given compound to be uniformly modified, and in fact more than one of the aforementioned modifications may be incorporated in a single compound or even at a single nucleoside within an oligonucleotide. The present invention also includes antisense compounds which are chimeric compounds. "Chimeric" antisense compounds or "chimeras," in the context of this invention, are antisense compounds, particularly oligonucleotides, which contain two or more chemically distinct regions, each made up of at least one monomer unit, i.e., a nucleotide in the case of an oligonucleotide compound. These oligonucleotides typically contain at least one region wherein the oligonucleotide is modified so as to confer upon the oligonucleotide increased resistance to nuclease degradation, increased cellular uptake, and/or increased binding affinity for the target nucleic acid. An additional region of the oligonucleotide may serve as a substrate for enzymes capable of cleaving RNA:DNA or RNA:RNA hybrids. By way of example, RNase H is a cellular endonuclease which cleaves the RNA strand of an RNA:DNA duplex. Activation of RNase H, therefore, results in cleavage of the RNA target, thereby greatly enhancing the efficiency of oligonucleotide inhibition of gene expression. Consequently, comparable results can often be obtained with shorter oligonucleotides when chimeric oligonucleotides are used, compared to phosphorothioate deoxyoligonucleotides hybridizing to the same target region. Chimeric antisense compounds of the invention may be formed as composite structures of two or more oligonucleotides, modified oligonucleotides, oligonucleosides and/or oligonucleotide mimetics as described above. Preferred chimeric antisense oligonucleotides incorporate at least one 2' modified sugar (preferably 2'-O-$(CH_2)_2$-O-CH_3) at the 3' terminal to confer nuclease resistance and a region with at least 4 contiguous 2'-H sugars to confer RNase H activity. Such compounds have also been referred to in the art as hybrids or gapmers. Preferred gapmers have a region of 2' modified sugars (preferably 2'-O-$(CH_2)_2$-O-CH_3) at the 3'-terminal and at the 5' terminal separated by at least one region having at least 4 contiguous 2'-H sugars and preferably incorporate phosphorothioate backbone linkages. Representative United States patents that teach the preparation of such hybrid structures include, but are not limited to, U.S. Pat. Nos. 5,013,830; 5,149,797; 5,220,007; 5,256,775; 5,366,878; 5,403,711; 5,491,133; 5,565,350; 5,623,065; 5,652,355; 5,652,356; and 5,700,922, each of which is herein incorporated by reference in its entirety.

The antisense compounds used in accordance with this invention may be conveniently and routinely made through the well-known technique of solid phase synthesis. Equipment for such synthesis is sold by several vendors including, for example, Applied Biosystems (Foster City, Calif.). Any other means for such synthesis known in the art may additionally or alternatively be employed. It is well known to use similar techniques to prepare oligonucleotides such as the phosphorothioates and alkylated derivatives. The compounds of the invention may also be admixed, encapsulated, conjugated or otherwise associated with other molecules, molecule structures or mixtures of compounds, as for example, liposomes, receptor targeted

molecules, oral, rectal, topical or other formulations, for assisting in uptake, distribution and/or absorption. Representative United States patents that teach the preparation of such uptake, distribution and/or absorption assisting formulations include, but are not limited to, U.S. Pat. Nos. 5,108,921; 5,354,844; 5,416,016; 5,459,127; 5,521,291; 5,543,158; 5,547,932; 5,583,020; 5,591,721; 4,426,330; 4,534,899; 5,013,556; 5,108,921; 5,213,804; 5,227,170; 5,264,221; 5,356,633; 5,395,619; 5,416,016; 5,417,978; 5,462,854; 5,469,854; 5,512,295; 5,527,528; 5,534,259; 5,543,152; 5,556,948; 5,580,575; and 5,595,756, each of which is herein incorporated by reference.

Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties, such as those described in WO 90/10048, and other moieties that increases affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents, such as ellipticine, and alkylating agents or metal complexes may be attached to sense or antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide sequence.

Antisense or sense oligonucleotides may be introduced into a cell containing the target nucleic acid sequence by any gene transfer method, including, for example, $CaPO_4$-mediated DNA transfection, electroporation, or by using gene transfer vectors such as Epstein-Barr virus. In a preferred procedure, an antisense or sense oligonucleotide is inserted into a suitable retroviral vector. A cell containing the target nucleic acid sequence is contacted with the recombinant retroviral vector, either *in vivo* or *ex vivo*. Suitable retroviral vectors include, but are not limited to, those derived from the murine retrovirus M-MuLV, N2 (a retrovirus derived from M-MuLV), or the double copy vectors designated DCT5A, DCT5B and DCT5C (see WO 90/13641).

Sense or antisense oligonucleotides also may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell.

Alternatively, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex, as described in WO 90/10448. The sense or antisense oligonucleotide-lipid complex is preferably dissociated within the cell by an endogenous lipase.

Antisense or sense RNA or DNA molecules are generally at least about 5 nucleotides in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750,

760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, or 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length.

Alternatively, a double stranded RNA can be generated. Double stranded RNA that are under 30 nucleotides in length will inhibit the expression of specific genes when introduced into a cell. This mechanism is known as RNA mediated interference (RNAi) and small (under 30 nuclotide) RNAs used as a reagent are known as siRNAs. TAT interfering RNAs may be identified and synthesized using known methods (Shi Y*Trends in Genetics* 19(1):9-12 (2003), WO/2003056012 and WO2003064621). siRNAs are useful to reduce the amount of gene expression in conditions where a reduction in the expression of the target gene would alleviate the condition or disorder.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related TAT coding sequences.

Nucleotide sequences encoding a TAT can also be used to construct hybridization probes for mapping the gene which encodes that TAT and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for TAT encode a protein which binds to another protein (example, where the TAT is a receptor), the TAT can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor TAT can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native TAT or a receptor for TAT. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode TAT or its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding TAT can be used to clone genomic DNA encoding TAT in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding TAT. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009.

Typically, particular cells would be targeted for TAT transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding TAT introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding TAT. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of TAT can be used to construct a TAT "knock out" animal which has a defective or altered gene encoding TAT as a result of homologous recombination between the endogenous gene encoding TAT and altered genomic DNA encoding TAT introduced into an embryonic stem cell of the animal. For example, cDNA encoding TAT can be used to clone genomic DNA encoding TAT in accordance with established techniques. A portion of the genomic DNA encoding TAT can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, Cell, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., Cell, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the TAT polypeptide.

Nucleic acid encoding the TAT polypeptides may also be used in gene therapy. In gene therapy applications, genes are introduced into cells in order to achieve *in vivo* synthesis of a therapeutically effective genetic product, for example for replacement of a defective gene. "Gene therapy" includes both conventional gene therapy where a lasting effect is achieved by a single treatment, and the administration of gene therapeutic agents, which involves the one time or repeated administration of a therapeutically effective DNA or mRNA. Antisense RNAs and DNAs can be used as therapeutic agents for blocking the expression of certain genes *in vivo*. It has already been shown that short antisense oligonucleotides can be imported into cells where they act as inhibitors, despite their low intracellular concentrations caused by their restricted uptake by the cell membrane. (Zamecnik *et al.*, Proc. Natl. Acad. Sci. USA 83:4143-4146 [1986]). The oligonucleotides can be modified to enhance their uptake, e.g. by substituting their negatively charged phosphodiester groups by uncharged groups.

There are a variety of techniques available for introducing nucleic acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells *in vitro*, or *in vivo* in the cells of

the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation, microinjection, cell fusion, DEAE-dextran, the calcium phosphate precipitation method, etc. The currently preferred *in vivo* gene transfer techniques include transfection with viral (typically retroviral) vectors and viral coat protein-liposome mediated transfection (Dzau et al., Trends in Biotechnology 11, 205-210 [1993]). In some situations it is desirable to provide the nucleic acid source with an agent that targets the target cells, such as an antibody specific for a cell surface membrane protein or the target cell, a ligand for a receptor on the target cell, etc. Where liposomes are employed, proteins which bind to a cell surface membrane protein associated with endocytosis may be used for targeting and/or to facilitate uptake, e.g. capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half-life. The technique of receptor-mediated endocytosis is described, for example, by Wu et al., J. Biol. Chem. 262, 4429-4432 (1987); and Wagner et al., Proc. Natl. Acad. Sci. USA 87, 3410-3414 (1990). For review of gene marking and gene therapy protocols see Anderson et al., Science 256, 808-813 (1992).

The nucleic acid molecules encoding the TAT polypeptides or fragments thereof described herein are useful for chromosome identification. In this regard, there exists an ongoing need to identify new chromosome markers, since relatively few chromosome marking reagents, based upon actual sequence data are presently available. Each TAT nucleic acid molecule of the present invention can be used as a chromosome marker.

The TAT polypeptides and nucleic acid molecules of the present invention may also be used diagnostically for tissue typing, wherein the TAT polypeptides of the present invention may be differentially expressed in one tissue as compared to another, preferably in a diseased tissue as compared to a normal tissue of the same tissue type. TAT nucleic acid molecules will find use for generating probes for PCR, Northern analysis, Southern analysis and Western analysis.

This invention encompasses methods of screening compounds to identify those that mimic the TAT polypeptide (agonists) or prevent the effect of the TAT polypeptide (antagonists). Screening assays for antagonist drug candidates are designed to identify compounds that bind or complex with the TAT polypeptides encoded by the genes identified herein, or otherwise interfere with the interaction of the encoded polypeptides with other cellular proteins, including e.g., inhibiting the expression of TAT polypeptide from cells. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates.

The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays, and cell-based assays, which are well characterized in the art.

All assays for antagonists are common in that they call for contacting the drug candidate with a TAT polypeptide encoded by a nucleic acid identified herein under conditions and for a time sufficient to allow these two components to interact.

In binding assays, the interaction is binding and the complex formed can be isolated or detected in the reaction mixture. In a particular embodiment, the TAT polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, e.g., on a microtiter plate, by covalent or non-covalent attachments.

Non-covalent attachment generally is accomplished by coating the solid surface with a solution of the TAT polypeptide and drying. Alternatively, an immobilized antibody, e.g., a monoclonal antibody, specific for the TAT polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, e.g., the coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, e.g., by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label immobilized on the surface indicates that complexing occurred. Where the originally non-immobilized component does not carry a label, complexing can be detected, for example, by using a labeled antibody specifically binding the immobilized complex.

If the candidate compound interacts with but does not bind to a particular TAT polypeptide encoded by a gene identified herein, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, e.g., cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers (Fields and Song, Nature (London), 340:245-246 (1989); Chien et al., Proc. Natl. Acad. Sci. USA, 88:9578-9582 (1991)) as disclosed by Chevray and Nathans, Proc. Natl. Acad. Sci. USA, 89: 5789-5793 (1991). Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, the other one functioning as the transcription-activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system") takes advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-*lacZ* reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for β-galactosidase. A complete kit (MATCHMAKER™) for identifying protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

Compounds that interfere with the interaction of a gene encoding a TAT polypeptide identified herein and other intra- or extracellular components can be tested as follows: usually a reaction mixture is prepared containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a candidate compound to inhibit binding, the reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation) between the test compound and the intra- or extracellular component present in the mixture is monitored as described hereinabove. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the test compound and its reaction partner.

To assay for antagonists, the TAT polypeptide may be added to a cell along with the compound to be screened for a particular activity and the ability of the compound to inhibit the activity of interest in the presence of the TAT polypeptide indicates that the compound is an antagonist to the TAT polypeptide. Alternatively, antagonists may be detected by combining the TAT polypeptide and a potential antagonist with membrane-bound TAT polypeptide receptors or recombinant receptors under appropriate conditions for a competitive inhibition assay. The TAT polypeptide can be labeled, such as by radioactivity, such that the number of TAT polypeptide molecules bound to the receptor can be used to determine the effectiveness of the potential antagonist. The gene encoding the receptor can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Coligan et al., Current Protocols in Immun., 1(2): Chapter 5 (1991). Preferably, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the TAT polypeptide and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the TAT polypeptide. Transfected cells that are grown on glass slides are exposed to labeled TAT polypeptide. The TAT polypeptide can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase. Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an interactive sub-pooling and re-screening process, eventually yielding a single clone that encodes the putative receptor.

As an alternative approach for receptor identification, labeled TAT polypeptide can be photoaffinity-linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE and exposed to X-ray film. The labeled complex containing the receptor can be excised, resolved into peptide fragments, and subjected to protein micro-sequencing. The amino acid sequence obtained from micro- sequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the gene encoding the putative receptor.

In another assay for antagonists, mammalian cells or a membrane preparation expressing the receptor would be incubated with labeled TAT polypeptide in the presence of the candidate compound. The ability of the compound to enhance or block this interaction could then be measured.

More specific examples of potential antagonists include an oligonucleotide that binds to the fusions of immunoglobulin with TAT polypeptide, and, in particular, antibodies including, without limitation, poly- and monoclonal antibodies and antibody fragments, single-chain antibodies, anti-idiotypic antibodies, and chimeric or humanized versions of such antibodies or fragments, as well as human antibodies and antibody fragments. Alternatively, a potential antagonist may be a closely related protein, for example, a mutated form of the TAT polypeptide that recognizes the receptor but imparts no effect, thereby competitively inhibiting the action of the TAT polypeptide.

Another potential TAT polypeptide antagonist is an antisense RNA or DNA construct prepared using antisense technology, where, e.g., an antisense RNA or DNA molecule acts to block directly the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are

based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes the mature TAT polypeptides herein, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res., 6:3073 (1979); Cooney et al., Science, 241: 456 (1988); Dervan et al., Science, 251:1360 (1991)), thereby preventing transcription and the production of the TAT polypeptide. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the TAT polypeptide (antisense - Okano, Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression (CRC Press: Boca Raton, FL, 1988). The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of the TAT polypeptide. When antisense DNA is used, oligodeoxyribonucleotides derived from the translation-initiation site, e.g., between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

Potential antagonists include small molecules that bind to the active site, the receptor binding site, or growth factor or other relevant binding site of the TAT polypeptide, thereby blocking the normal biological activity of the TAT polypeptide. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules, preferably soluble peptides, and synthetic non-peptidyl organic or inorganic compounds.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques. For further details see, e.g., Rossi, Current Biology, 4:469-471 (1994), and PCT publication No. WO 97/33551 (published September 18, 1997).

Nucleic acid molecules in triple-helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple-helix formation via Hoogsteen base-pairing rules, which generally require sizeable stretches of purines or pyrimidines on one strand of a duplex. For further details see, e.g., PCT publication No. WO 97/33551, *supra*.

These small molecules can be identified by any one or more of the screening assays discussed hereinabove and/or by any other screening techniques well known for those skilled in the art.

Isolated TAT polypeptide-encoding nucleic acid can be used herein for recombinantly producing TAT polypeptide using techniques well known in the art and as described herein. In turn, the produced TAT polypeptides can be employed for generating anti-TAT antibodies using techniques well known in the art and as described herein.

Antibodies specifically binding a TAT polypeptide identified herein, as well as other molecules identified by the screening assays disclosed hereinbefore, can be administered for the treatment of various disorders, including cancer, in the form of pharmaceutical compositions.

If the TAT polypeptide is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, lipofections or liposomes can also be used to deliver the antibody, or an antibody

fragment, into cells. Where antibody fragments are used, the smallest inhibitory fragment that specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody, peptide molecules can be designed that retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. See, *e.g.,* Marasco *et al.,* Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993).

The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Manassas, VA.

EXAMPLE 1: Tissue Expression Profiling Using GeneExpress®

A proprietary database containing gene expression information (GeneExpress®, Gene Logic Inc., Gaithersburg, MD) was analyzed in an attempt to identify polypeptides (and their encoding nucleic acids) whose expression is significantly upregulated in a particular tumor tissue(s) of interest as compared to other tumor(s) and/or normal tissues. Specifically, analysis of the GeneExpress® database was conducted using either software available through Gene Logic Inc., Gaithersburg, MD, for use with the GeneExpress® database or with proprietary software written and developed at Genentech, Inc. for use with the GeneExpress® database. The rating of positive hits in the analysis is based upon several criteria including, for example, tissue specificity, tumor specificity and expression level in normal essential and/or normal proliferating tissues. The following is a list of molecules whose tissue expression profile as determined from an analysis of the GeneExpress® database evidences high tissue expression and significant upregulation of expression in a specific tumor or tumors as compared to other tumor(s) and/or normal tissues and optionally relatively low expression in normal essential and/or normal proliferating tissues. As such, the molecules listed below are excellent polypeptide targets for the diagnosis and therapy of cancer in mammals.

Molecule	upregulation of expression in:	as compared to:

DNA77507 (TAT161)	breast tumor	normal breast tissue
DNA77507 (TAT161)	colon tumor	normal colon tissue
DNA77507 (TAT161)	lung tumor	normal lung tissue
DNA77507 (TAT161)	kidney tumor	normal kidney tissue
DNA77507 (TAT161)	liver tumor	normal liver tissue
DNA77507 (TAT161)	ovarian tumor	normal ovarian tissue
DNA77507 (TAT161)	pancreatic tumor	normal pancreatic tissue
DNA77507 (TAT161)	rectum tumor	normal rectum tissue
DNA77507 (TAT161)	skin tumor	normal skin tissue
DNA77507 (TAT161)	uterine tumor	normal uterine tissue
DNA77507 (TAT161)	brain tumor	normal brain tissue
DNA77507 (TAT161)	soft tissue tumor	normal soft tissue
DNA77507 (TAT161)	bone tumor	normal bone tissue
DNA80894 (TAT101)	breast tumor	normal breast tissue
DNA82343 (TAT157)	colon tumor	normal colon tissue
DNA82343 (TAT157)	ovarian tumor	normal ovarian tissue
DNA82343 (TAT157)	stomach tumor	normal stomach tissue
DNA82343 (TAT157)	liver tumor	normal liver tissue
DNA82343 (TAT157)	rectum tumor	normal rectum tissue
DNA82343 (TAT157)	small intestine tumor	normal small intestine tissue
DNA82343 (TAT157)	esophagus tumor	normal esophagus tissue
DNA82343 (TAT157)	testis tumor	normal testis tissue
DNA82343 (TAT157)	thymus tumor	normal thymus tissue
DNA87994 (TAT160)	breast tumor	normal breast tissue
DNA87994 (TAT160)	pancreatic tumor	normal pancreatic tissue
DNA87994 (TAT160)	rectum tumor	normal rectum tissue
DNA87994 (TAT160)	colon tumor	normal colon tissue
DNA87994 (TAT160)	esophagus tumor	normal esophagus tissue
DNA87994 (TAT160)	ovarian tumor	normal ovarian tissue
DNA87994 (TAT160)	lung tumor	normal lung tissue
DNA87994 (TAT160)	uterine tumor	normal uterine tissue
DNA88131 (TAT158)	bone tumor	normal bone tissue
DNA88131 (TAT158)	breast tumor	normal breast tissue
DNA88131 (TAT158)	colon tumor	normal colon tissue
DNA88131 (TAT158)	uterine tumor	normal uterine tissue
DNA88131 (TAT158)	esophagus tumor	normal esophagus tissue
DNA88131 (TAT158)	lung tumor	normal lung tissue

DNA88131 (TAT158)	ovarian tumor	normal ovarian tissue
DNA88131 (TAT158)	pancreatic tumor	normal pancreatic tissue
DNA88131 (TAT158)	prostate tumor	normal prostate tissue
DNA88131 (TAT158)	skin tumor	normal skin tissue
DNA88131 (TAT158)	soft tissue tumor	normal soft tissue
DNA88131 (TAT158)	stomach tumor	normal stomach tissue

Molecule	upregulation of expression in:	as compared to:
DNA88131 (TAT158)	rectum tumor	normal rectum tissue
DNA88131 (TAT158)	neuroendocrine tumor	normal neuroendocrine tissue
DNA88131 (TAT158)	brain tumor	normal brain tissue
DNA95930 (TAT110)	colon tumor	normal colon tissue
DNA95930 (TAT110)	uterine tumor	normal uterine tissue
DNA95930 (TAT110)	endometrial tumor	normal endometrial tissue
DNA95930 (TAT110)	rectum tumor	normal rectum tissue
DNA95930 (TAT110)	ovarian tumor	normal ovarian tissue
DNA95930 (TAT110)	breast tumor	normal breast tissue
DNA95930 (TAT110)	lung tumor	normal lung tissue
DNA95930 (TAT110)	prostate tumor	normal prostate tissue
DNA95930-1 (TAT210)	colon tumor	normal colon tissue
DNA95930-1 (TAT210)	uterine tumor	normal uterine tissue
DNA95930-1 (TAT210)	endometrial tumor	normal endometrial tissue
DNA95930-1 (TAT210)	rectum tumor	normal rectum tissue
DNA95930-1 (TAT210)	ovarian tumor	normal ovarian tissue
DNA95930-1 (TAT210)	breast tumor	normal breast tissue
DNA95930-1 (TAT210)	lung tumor	normal lung tissue
DNA95930-1 (TAT210)	prostate tumor	normal prostate tissue
DNA96917 (TAT159)	pancreatic tumor	normal pancreatic tissue
DNA96917 (TAT159)	lung tumor	normal lung tissue
DNA96917 (TAT159)	liver tumor	normal liver tissue
DNA96930 (TAT112)	breast tumor	normal breast tissue
DNA96930 (TAT112)	colon tumor	normal colon tissue
DNA96930 (TAT112)	rectum tumor	normal rectum tissue
DNA96930 (TAT112)	uterine tumor	normal uterine tissue
DNA96930 (TAT112)	lung tumor	normal lung tissue
DNA96930 (TAT112)	ovarian tumor	normal ovarian tissue
DNA96930 (TAT112)	pancreatic tumor	normal pancreatic tissue

DNA96930 (TAT112)	stomach tumor	normal stomach tissue
DNA96936 (TAT147)	breast tumor	normal breast tissue
DNA96936 (TAT147)	colon tumor	normal colon tissue
DNA96936 (TAT147)	testis tumor	normal testis tissue
DNA96936 (TAT147)	ovarian tumor	normal ovarian tissue
DNA98565 (TAT145)	brain tumor	normal brain tissue
DNA98565 (TAT145)	glioma	normal glial tissue
DNA246435 (TAT152)	brain tumor	normal brain tissue
DNA246435 (TAT152)	glioma	normal glial tissue
DNA98591 (TAT162)	colon tumor	normal colon tissue
DNA98591 (TAT162)	rectum tumor	normal rectum tissue
DNA98591 (TAT162)	ovarian tumor	normal ovarian tissue
DNA98591 (TAT162)	pancreatic tumor	normal pancreatic tissue
DNA98591 (TAT162)	stomach tumor	normal stomach tissue
DNA108809 (TAT114)	colon tumor	normal colon tissue
DNA108809 (TAT114)	kidney tumor	normal kidney tissue
DNA119488 (TAT119)	colon tumor	normal colon tissue
DNA119488 (TAT119)	lung tumor	normal lung tissue
DNA119488 (TAT119)	rectum tumor	normal rectum tissue
DNA143493 (TAT103)	breast tumor	normal breast tissue
DNA167234 (TAT130)	prostate tumor	normal prostate tissue
DNA235621 (TAT166)	prostate tumor	normal prostate tissue
DNA235621 (TAT166)	liver tumor	normal liver tissue
DNA176766 (TAT132)	kidney tumor	normal kidney tissue
DNA176766 (TAT132)	ovarian tumor	normal ovarian tissue
DNA176766 (TAT132)	uterine tumor	normal uterine tissue

Molecule	upregulation of expression in:	as compared to:
DNA236463 (TAT150)	kidney tumor	normal kidney tissue
DNA236463 (TAT150)	ovarian tumor	normal ovarian tissue
DNA236463 (TAT150)	uterine tumor	normal uterine tissue
DNA181162 (TAT129)	prostate tumor	normal prostate tissue
DNA188221 (TAT111)	colon tumor	normal colon tissue
DNA188221 (TAT111)	endometrial tumor	normal endometrial tissue
DNA188221 (TAT111)	stomach tumor	normal stomach tissue
DNA233876 (TAT146)	colon tumor	normal colon tissue
DNA233876 (TAT146)	endometrial tumor	normal endometrial tissue

DNA233876 (TAT146)	stomach tumor	normal stomach tissue
DNA193891 (TAT148)	colon tumor	normal colon tissue
DNA248170 (TAT187)	colon tumor	normal colon tissue
DNA248170 (TAT187)	breast tumor	normal breast tissue
DNA194628 (TAT118)	kidney tumor	normal kidney tissue
DNA246415 (TAT167)	kidney tumor	normal kidney tissue
DNA215609 (TAT113)	colon tumor	normal colon tissue
DNA215609 (TAT113)	rectum tumor	normal rectum tissue
DNA220432 (TAT128)	prostate tumor	normal prostate tissue
DNA226094 (TAT164)	breast tumor	normal breast tissue
DNA226094 (TAT164)	brain tumor	normal brain tissue
DNA226094 (TAT164)	lung tumor	normal lung tissue
DNA226094 (TAT164)	skin tumor	normal skin tissue
DNA226165 (TAT122)	breast tumor	normal breast tissue
DNA226165 (TAT122)	endometrial tumor	normal endometrial tissue
DNA226165 (TAT122)	kidney tumor	normal kidney tissue
DNA226165 (TAT122)	lung tumor	normal lung tissue
DNA226165 (TAT122)	ovarian tumor	normal ovarian tissue
DNA226165 (TAT122)	colon tumor	normal colon tissue
DNA226165 (TAT122)	rectum tumor	normal rectum tissue
DNA226165 (TAT122)	skin tumor	normal skin tissue
DNA226165 (TAT122)	soft tissue tumor	normal soft tissue tissue
DNA226165 (TAT122)	bladder tumor	normal bladder tissue
DNA226237 (TAT117)	kidney tumor	normal kidney tissue
DNA246450 (TAT168)	kidney tumor	normal kidney tissue
DNA226456 (TAT144)	breast tumor	normal breast tissue
DNA226456 (TAT144)	colon tumor	normal colon tissue
DNA226456 (TAT144)	rectum tumor	normal rectum tissue
DNA226456 (TAT144)	endometrial tumor	normal endometrial tissue
DNA226456 (TAT144)	kidney tumor	normal kidney tissue
DNA226456 (TAT144)	lung tumor	normal lung tissue
DNA226456 (TAT144)	ovarian tumor	normal ovarian tissue
DNA226456 (TAT144)	skin tumor	normal skin tissue
DNA237637 (TAT188)	breast tumor	normal breast tissue
DNA237637 (TAT188)	colon tumor	normal colon tissue
DNA237637 (TAT188)	rectum tumor	normal rectum tissue
DNA237637 (TAT188)	endometrial tumor	normal endometrial tissue

DNA237637 (TAT188)	kidney tumor	normal kidney tissue
DNA237637 (TAT188)	lung tumor	normal lung tissue
DNA237637 (TAT188)	ovarian tumor	normal ovarian tissue
DNA237637 (TAT188)	skin tumor	normal skin tissue
DNA237637 (TAT188)	liver tumor	normal liver tissue
DNA237637 (TAT188)	lung tumor	normal lung tissue
DNA226539 (TAT126)	breast tumor	normal breast tissue
DNA226539 (TAT126)	colon tumor	normal colon tissue
DNA226539 (TAT126)	rectum tumor	normal rectum tissue

Molecule	upregulation of expression in:	as compared to:
DNA226539 (TAT126)	endometrial tumor	normal endometrial tissue
DNA226539 (TAT126)	lung tumor	normal lung tissue
DNA226539 (TAT126)	ovarian tumor	normal ovarian tissue
DNA226539 (TAT126)	pancreatic tumor	normal pancreatic tissue
DNA236511 (TAT151)	breast tumor	normal breast tissue
DNA236511 (TAT151)	colon tumor	normal colon tissue
DNA236511 (TAT151)	rectum tumor	normal rectum tissue
DNA236511 (TAT151)	endometrial tumor	normal endometrial tissue
DNA236511 (TAT151)	lung tumor	normal lung tissue
DNA236511 (TAT151)	ovarian tumor	normal ovarian tissue
DNA236511 (TAT151)	pancreatic tumor	normal pancreatic tissue
DNA226771 (TAT115)	breast tumor	normal breast tissue
DNA226771 (TAT115)	colon tumor	normal colon tissue
DNA227087 (TAT163)	breast tumor	normal breast tissue
DNA227087 (TAT163)	colon tumor	normal colon tissue
DNA227087 (TAT163)	rectum tumor	normal rectum tissue
DNA227087 (TAT163)	lung tumor	normal lung tissue
DNA227087 (TAT163)	ovarian tumor	normal ovarian tissue
DNA227087 (TAT163)	prostate tumor	normal prostate tissue
DNA227087 (TAT163)	endocrine tumor	normal endocrine tissue
DNA227087 (TAT163)	kidney tumor	normal kidney tissue
DNA227087 (TAT163)	liver tumor	normal liver tissue
DNA227087 (TAT163)	nervous system tumor	normal nervous system tissue
DNA227087 (TAT163)	pancreatic tumor	normal pancreatic tissue
DNA227087 (TAT163)	uterine tumor	normal uterine tissue
DNA227087 (TAT163)	small intestine tumor	normal small intestine tissue

DNA227087 (TAT163)	lymphoid tumor	normal lymphoid tissue
DNA266307 (TAT227)	breast tumor	normal breast tissue
DNA266307 (TAT227)	colon tumor	normal colon tissue
DNA266307 (TAT227)	rectum tumor	normal rectum tissue
DNA266307 (TAT227)	lung tumor	normal lung tissue
DNA266307 (TAT227)	ovarian tumor	normal ovarian tissue
DNA266307 (TAT227)	prostate tumor	normal prostate tissue
DNA266307 (TAT227)	endocrine tumor	normal endocrine tissue
DNA266307 (TAT227)	kidney tumor	normal kidney tissue
DNA266307 (TAT227)	liver tumor	normal liver tissue
DNA266307 (TAT227)	nervous system tumor	normal nervous system tissue
DNA266307 (TAT227)	pancreatic tumor	normal pancreatic tissue
DNA266307 (TAT227)	uterine tumor	normal uterine tissue
DNA266307 (TAT227)	small intestine tumor	normal small intestine tissue
DNA266307 (TAT227)	lymphoid tumor	normal lymphoid tissue
DNA266311 (TAT228)	breast tumor	normal breast tissue
DNA266311 (TAT228)	colon tumor	normal colon tissue
DNA266311 (TAT228)	rectum tumor	normal rectum tissue
DNA266311 (TAT228)	lung tumor	normal lung tissue
DNA266311 (TAT228)	ovarian tumor	normal ovarian tissue
DNA266311 (TAT228)	prostate tumor	normal prostate tissue
DNA266311 (TAT228)	endocrine tumor	normal endocrine tissue
DNA266311 (TAT228)	kidney tumor	normal kidney tissue
DNA266311 (TAT228)	liver tumor	normal liver tissue
DNA266311 (TAT228)	nervous system tumor	normal nervous system tissue
DNA266311 (TAT228)	pancreatic tumor	normal pancreatic tissue
DNA266311 (TAT228)	uterine tumor	normal uterine tissue
DNA266311 (TAT228)	small intestine tumor	normal small intestine tissue
DNA266311 (TAT228)	lymphoid tumor	normal lymphoid tissue

Molecule	upregulation of expression in:	as compared to:
DNA266312 (TAT229)	breast tumor	normal breast tissue
DNA266312 (TAT229)	colon tumor	normal colon tissue
DNA266312 (TAT229)	rectum tumor	normal rectum tissue
DNA266312 (TAT229)	lung tumor	normal lung tissue
DNA266312 (TAT229)	ovarian tumor	normal ovarian tissue
DNA266312 (TAT229)	prostate tumor	normal prostate tissue

DNA266312 (TAT229)	endocrine tumor	normal endocrine tissue
DNA266312 (TAT229)	kidney tumor	normal kidney tissue
DNA266312 (TAT229)	liver tumor	normal liver tissue
DNA266312 (TAT229)	nervous system tumor	normal nervous system tissue
DNA266312 (TAT229)	pancreatic tumor	normal pancreatic tissue
DNA266312 (TAT229)	uterine tumor	normal uterine tissue
DNA266312 (TAT229)	small intestine tumor	normal small intestine tissue
DNA266312 (TAT229)	lymphoid tumor	normal lymphoid tissue
DNA266313 (TAT230)	breast tumor	normal breast tissue
DNA266313 (TAT230)	colon tumor	normal colon tissue
DNA266313 (TAT230)	rectum tumor	normal rectum tissue
DNA266313 (TAT230)	lung tumor	normal lung tissue
DNA266313 (TAT230)	ovarian tumor	normal ovarian tissue
DNA266313 (TAT230)	prostate tumor	normal prostate tissue
DNA266313 (TAT230)	endocrine tumor	normal endocrine tissue
DNA266313 (TAT230)	kidney tumor	normal kidney tissue
DNA266313 (TAT230)	liver tumor	normal liver tissue
DNA266313 (TAT230)	nervous system tumor	normal nervous system tissue
DNA266313 (TAT230)	pancreatic tumor	normal pancreatic tissue
DNA266313 (TAT230)	uterine tumor	normal uterine tissue
DNA266313 (TAT230)	small intestine tumor	normal small intestine tissue
DNA266313 (TAT230)	lymphoid tumor	normal lymphoid tissue
DNA227224 (TAT121)	breast tumor	normal breast tissue
DNA227224 (TAT121)	colon tumor	normal colon tissue
DNA227224 (TAT121)	rectum tumor	normal rectum tissue
DNA227224 (TAT121)	endometrial tumor	normal endometrial tissue
DNA227224 (TAT121)	kidney tumor	normal kidney tissue
DNA227224 (TAT121)	lung tumor	normal lung tissue
DNA227224 (TAT121)	ovarian tumor	normal ovarian tissue
DNA227224 (TAT121)	skin tumor	normal skin tissue
DNA227224 (TAT121)	testis tumor	normal testis tissue
DNA227224 (TAT121)	bladder tumor	normal bladder tissue
DNA247486 (TAT183)	breast tumor	normal breast tissue
DNA247486 (TAT183)	colon tumor	normal colon tissue
DNA247486 (TAT183)	rectum tumor	normal rectum tissue
DNA247486 (TAT183)	endometrial tumor	normal endometrial tissue
DNA247486 (TAT183)	kidney tumor	normal kidney tissue

DNA247486 (TAT183)	lung tumor	normal lung tissue
DNA247486 (TAT183)	ovarian tumor	normal ovarian tissue
DNA247486 (TAT183)	skin tumor	normal skin tissue
DNA247486 (TAT183)	testis tumor	normal testis tissue
DNA247486 (TAT183)	bladder tumor	normal bladder tissue
DNA227800 (TAT131)	prostate tumor	normal prostate tissue
DNA228199 (TAT127)	breast tumor	normal breast tissue
DNA228199 (TAT127)	endometrial tumor	normal endometrial tissue
DNA228199 (TAT127)	ovarian tumor	normal ovarian tissue
DNA228199 (TAT127)	pancreatic tumor	normal pancreatic tissue
DNA228199 (TAT127)	lung tumor	normal lung tissue
DNA228201 (TAT116)	colon tumor	normal colon tissue

Molecule	upregulation of expression in:	as compared to:
DNA228201 (TAT116)	rectum tumor	normal rectum tissue
DNA247488 (TAT189)	colon tumor	normal colon tissue
DNA247488 (TAT189)	rectum tumor	normal rectum tissue
DNA236538 (TAT190)	colon tumor	normal colon tissue
DNA236538 (TAT190)	rectum tumor	normal rectum tissue
DNA247489 (TAT191)	colon tumor	normal colon tissue
DNA247489 (TAT191)	rectum tumor	normal rectum tissue
DNA228211 (TAT133)	uterine tumor	normal uterine tissue
DNA233937 (TAT186)	uterine tumor	normal uterine tissue
DNA233937 (TAT186)	ovarian tumor	normal ovarian tissue
DNA228994 (TAT124)	lung tumor	normal lung tissue
DNA228994 (TAT124)	ovarian tumor	normal ovarian tissue
DNA228994 (TAT124)	skin tumor	normal skin tissue
DNA228994 (TAT124)	breast tumor	normal breast tissue
DNA229410 (TAT105)	breast tumor	normal breast tissue
DNA229411 (TAT107)	breast tumor	normal breast tissue
DNA229413 (TAT108)	breast tumor	normal breast tissue
DNA229700 (TAT139)	breast tumor	normal breast tissue
DNA231312 (TAT143)	breast tumor	normal breast tissue
DNA231312 (TAT143)	colon tumor	normal colon tissue
DNA231542 (TAT100)	brain tumor	normal brain tissue
DNA231542 (TAT100)	glioma	normal glial tissue
DNA231542-1 (TAT284)	brain tumor	normal brain tissue

DNA231542-1 (TAT284)	glioma	normal glial tissue
DNA231542-2 (TAT285)	brain tumor	normal brain tissue
DNA231542-2 (TAT285)	glioma	normal glial tissue
DNA297393 (TAT285-1)	brain tumor	normal brain tissue
DNA297393 (TAT285-1)	glioma	normal glial tissue
DNA234833 (TAT149)	colon tumor	normal colon tissue
DNA268022 (TAT231)	colon tumor	normal colon tissue
DNA268022 (TAT231)	breast tumor	normal breast tissue
DNA268022 (TAT231)	ovarian tumor	normal ovarian tissue
DNA236246 (TAT153)	breast tumor	normal breast tissue
DNA236343 (TAT104)	breast tumor	normal breast tissue
DNA236493 (TAT141)	breast tumor	normal breast tissue
DNA236493 (TAT141)	glioblastoma tumor	normal glial tissue
DNA236534 (TAT102)	breast tumor	normal breast tissue
DNA236534 (TAT102)	colon tumor	normal colon tissue
DNA236534 (TAT102)	rectum tumor	normal rectum tissue
DNA236534 (TAT102)	cervical tumor	normal cervical tissue
DNA236534 (TAT102)	endometrial tumor	normal endometrial tissue
DNA236534 (TAT102)	lung tumor	normal lung tissue
DNA236534 (TAT102)	ovarian tumor	normal ovarian tissue
DNA236534 (TAT102)	pancreatic tumor	normal pancreatic tissue
DNA236534 (TAT102)	prostate tumor	normal prostate tissue
DNA236534 (TAT102)	stomach tumor	normal stomach tissue
DNA236534 (TAT102)	bladder tumor	normal bladder tissue
DNA246430 (TAT109)	breast tumor	normal breast tissue
DNA246430 (TAT109)	prostate tumor	normal prostate tissue
DNA247480 (TAT142)	breast tumor	normal breast tissue
DNA247480 (TAT142)	lung tumor	normal lung tissue
DNA264454 (TAT106)	breast tumor	normal breast tissue

EXAMPLE 2: Microarray Analysis to Detect Upregulation of TAT Polypeptides in Cancerous Tumors

Nucleic acid microarrays, often containing thousands of gene sequences, are useful for identifying differentially expressed genes in diseased tissues as compared to their normal counterparts. Using nucleic acid microarrays, test and control mRNA samples from test and control tissue samples are reverse transcribed and labeled to generate cDNA probes. The cDNA probes are then hybridized to an array of nucleic acids immobilized on a solid support. The array is configured such that the sequence and position of each member of the array is

known. For example, a selection of genes known to be expressed in certain disease states may be arrayed on a solid support. Hybridization of a labeled probe with a particular array member indicates that the sample from which the probe was derived expresses that gene. If the hybridization signal of a probe from a test (disease tissue) sample is greater than hybridization signal of a probe from a control (normal tissue) sample, the gene or genes overexpressed in the disease tissue are identified. The implication of this result is that an overexpressed protein in a diseased tissue is useful not only as a diagnostic marker for the presence of the disease condition, but also as a therapeutic target for treatment of the disease condition.

The methodology of hybridization of nucleic acids and microarray technology is well known in the art. In one example, the specific preparation of nucleic acids for hybridization and probes, slides, and hybridization conditions are all detailed in PCT Patent Application Serial No. PCT/US01/10482, filed on March 30, 2001 and which is herein incorporated by reference.

In the present example, cancerous tumors derived from various human tissues were studied for upregulated gene expression relative to cancerous tumors from different tissue types and/or non-cancerous human tissues in an attempt to identify those polypeptides which are overexpressed in a particular cancerous tumor(s). In certain experiments, cancerous human tumor tissue and non-cancerous human tumor tissue of the same tissue type (often from the same patient) were obtained and analyzed for TAT polypeptide expression. Additionally, cancerous human tumor tissue from any of a variety of different human tumors was obtained and compared to a "universal" epithelial control sample which was prepared by pooling non-cancerous human tissues of epithelial origin, including liver, kidney, and lung. mRNA isolated from the pooled tissues represents a mixture of expressed gene products from these different tissues. Microarray hybridization experiments using the pooled control samples generated a linear plot in a 2-color analysis. The slope of the line generated in a 2-color analysis was then used to normalize the ratios of (test:control detection) within each experiment. The normalized ratios from various experiments were then compared and used to identify clustering of gene expression. Thus, the pooled "universal control" sample not only allowed effective relative gene expression determinations in a simple 2-sample comparison, it also allowed multi-sample comparisons across several experiments.

In the present experiments, nucleic acid probes derived from the herein described TAT polypeptide-encoding nucleic acid sequences were used in the creation of the microarray and RNA from various tumor tissues were used for the hybridization thereto. Below is shown the results of these experiments, demonstrating that various TAT polypeptides of the present invention are significantly overexpressed in various human tumor tissues as compared to their normal counterpart tissue(s). Moreover, all of the molecules shown below are significantly overexpressed in their specific tumor tissue(s) as compared to in the "universal" epithelial control. As described above, these data demonstrate that the TAT polypeptides of the present invention are useful not only as diagnostic markers for the presence of one or more cancerous tumors, but also serve as therapeutic targets for the treatment of those tumors.

Molecule	upregulation of expression in:	as compared to:
DNA95930 (TAT110)	colon tumor	normal colon tissue
DNA95930 (TAT110)	lung tumor	normal lung tissue

DNA95930 (TAT110)	prostate tumor	normal prostate tissue
DNA95930 (TAT110)	endometrial tumor	normal endometrial tissue
DNA95930 (TAT110)	ovarian tumor	normal ovarian tissue
DNA95930-1 (TAT210)	colon tumor	normal colon tissue
DNA95930-1 (TAT210)	lung tumor	normal lung tissue
DNA95930-1 (TAT210)	prostate tumor	normal prostate tissue
DNA95930-1 (TAT210)	endometrial tumor	normal endometrial tissue
DNA95930-1 (TAT210)	ovarian tumor	normal ovarian tissue
DNA96930 (TAT112)	colon tumor	normal colon tissue
DNA96930 (TAT112)	breast tumor	normal breast tissue
DNA96930 (TAT112)	lung tumor	normal lung tissue
DNA96936 (TAT147)	breast tumor	normal breast tissue
DNA96936 (TAT147)	colon tumor	normal colon tissue
DNA96936 (TAT147)	ovarian tumor	normal ovarian tissue
DNA96936 (TAT147)	prostate tumor	normal prostate tissue
DNA108809 (TAT114)	colon tumor	normal colon tissue
DNA119488 (TAT119)	colon tumor	normal colon tissue
DNA119488 (TAT119)	lung tumor	normal lung tissue
DNA143493 (TAT103)	breast tumor	normal breast tissue
DNA181162 (TAT129)	prostate tumor	normal prostate tissue
DNA188221 (TAT111)	colon tumor	normal colon tissue
DNA188221 (TAT111)	lung tumor	normal lung tissue
DNA188221 (TAT111)	ovarian tumor	normal ovarian tissue
DNA233876 (TAT146)	colon tumor	normal colon tissue
DNA233876 (TAT146)	lung tumor	normal lung tissue
DNA233876 (TAT146)	ovarian tumor	normal ovarian tissue
DNA210499 (TAT123)	ovarian tumor	normal ovarian tissue
DNA210499 (TAT123)	lung tumor	normal lung tissue
DNA219894 (TAT211)	ovarian tumor	normal ovarian tissue
DNA219894 (TAT211)	lung tumor	normal lung tissue
DNA215609 (TAT113)	colon tumor	normal colon tissue
DNA220432 (TAT128)	prostate tumor	normal prostate tissue
DNA226165 (TAT122)	breast tumor	normal breast tissue
DNA226165 (TAT122)	colon tumor	normal colon tissue
DNA226165 (TAT122)	rectum tumor	normal rectum tissue
DNA226165 (TAT122)	lung tumor	normal lung tissue
DNA226165 (TAT122)	ovarian tumor	normal ovarian tissue

DNA226165 (TAT122)	prostate tumor	normal prostate tissue
DNA226456 (TAT144)	breast tumor	normal breast tissue
DNA226456 (TAT144)	colon tumor	normal colon tissue
DNA237637 (TAT188)	breast tumor	normal breast tissue
DNA237637 (TAT188)	colon tumor	normal colon tissue
DNA226539 (TAT126)	rectum tumor	normal rectum tissue
DNA226539 (TAT126)	colon tumor	normal colon tissue
DNA226539 (TAT126)	lung tumor	normal lung tissue
DNA226539 (TAT126)	ovarian tumor	normal ovarian tissue
DNA236511 (TAT151)	rectum tumor	normal rectum tissue
DNA236511 (TAT151)	colon tumor	normal colon tissue
DNA236511 (TAT151)	lung tumor	normal lung tissue

Molecule	upregulation of expression in:	as compared to:
DNA236511 (TAT151)	ovarian tumor	normal ovarian tissue
DNA226771 (TAT115)	colon tumor	normal colon tissue
DNA227224 (TAT121)	ovarian tumor	normal ovarian tissue
DNA227224 (TAT121)	rectum tumor	normal rectum tissue
DNA227224 (TAT121)	colon tumor	normal colon tissue
DNA227224 (TAT121)	lung tumor	normal lung tissue
DNA227224 (TAT121)	breast tumor	normal breast tissue
DNA227224 (TAT121)	prostate tumor	normal prostate tissue
DNA247486 (TAT183)	ovarian tumor	normal ovarian tissue
DNA247486 (TAT183)	rectum tumor	normal rectum tissue
DNA247486 (TAT183)	colon tumor	normal colon tissue
DNA247486 (TAT183)	lung tumor	normal lung tissue
DNA247486 (TAT183)	breast tumor	normal breast tissue
DNA247486 (TAT183)	prostate tumor	normal prostate tissue
DNA228199 (TAT127)	ovarian tumor	normal ovarian tissue
DNA228199 (TAT127)	lung tumor	normal lung tissue
DNA228201 (TAT116)	colon tumor	normal colon tissue
DNA247488 (TAT189)	colon tumor	normal colon tissue
DNA236538 (TAT190)	colon tumor	normal colon tissue
DNA247489 (TAT191)	colon tumor	normal colon tissue
DNA228994 (TAT124)	lung tumor	normal lung tissue
DNA228994 (TAT124)	breast tumor	normal breast tissue
DNA228994 (TAT124)	ovarian tumor	normal ovarian tissue

DNA231312 (TAT143)	colon tumor	normal colon tissue
DNA231542 (TAT100)	brain tumor	normal brain tissue
DNA231542 (TAT100)	glioma	normal glial tissue
DNA231542-1 (TAT284)	brain tumor	normal brain tissue
DNA231542-1 (TAT284)	glioma	normal glial tissue
DNA231542-2 (TAT285)	brain tumor	normal brain tissue
DNA231542-2 (TAT285)	glioma	normal glial tissue
DNA297393 (TAT285-1)	brain tumor	normal brain tissue
DNA297393 (TAT285-1)	glioma	normal glial tissue
DNA236246 (TAT153)	breast tumor	normal breast tissue
DNA236343 (TAT104)	breast tumor	normal breast tissue
DNA236534 (TAT102)	breast tumor	normal breast tissue
DNA236534 (TAT102)	colon tumor	normal colon tissue
DNA246430 (TAT109)	prostate tumor	normal prostate tissue
DNA264454 (TAT106)	breast tumor	normal breast tissue
DNA98565 (TAT145)	glioma	normal brain tissue
DNA246435 (TAT152)	glioma	normal brain tissue
DNA226094 (TAT164)	glioma	normal brain tissue

EXAMPLE 3: Quantitative Analysis of TAT mRNA Expression

In this assay, a 5' nuclease assay (for example, TaqMan®) and real-time quantitative PCR (for example, ABI Prizm 7700 Sequence Detection System® (Perkin Elmer, Applied Biosystems Division, Foster City, CA)), were used to find genes that are significantly overexpressed in a cancerous tumor or tumors as compared to other cancerous tumors or normal non-cancerous tissue. The 5' nuclease assay reaction is a fluorescent PCR-based technique which makes use of the 5' exonuclease activity of Taq DNA polymerase enzyme to monitor gene expression in real time. Two oligonucleotide primers (whose sequences are based upon the gene or EST sequence of interest) are used to generate an amplicon typical of a PCR reaction. A third oligonucleotide, or probe, is designed to detect nucleotide sequence located between the two PCR primers. The probe is non-extendible by Taq DNA polymerase enzyme, and is labeled with a reporter fluorescent dye and a quencher fluorescent dye. Any laser-induced emission from the reporter dye is quenched by the quenching dye when the two dyes are located close together as they are on the probe. During the PCR amplification reaction, the Taq DNA polymerase enzyme cleaves the probe in a template-dependent manner. The resultant probe fragments disassociate in solution, and signal from the released reporter dye is free from the quenching effect of the second fluorophore. One molecule of reporter dye is liberated for each new molecule synthesized, and detection of the unquenched reporter dye provides the basis for quantitative interpretation of the data.

The 5' nuclease procedure is run on a real-time quantitative PCR device such as the ABI Prism 7700TM Sequence Detection. The system consists of a thermocycler, laser, charge-coupled device (CCD) camera and

computer. The system amplifies samples in a 96-well format on a thermocycler. During amplification, laser-induced fluorescent signal is collected in real-time through fiber optics cables for all 96 wells, and detected at the CCD. The system includes software for running the instrument and for analyzing the data.

The starting material for the screen was mRNA isolated from a variety of different cancerous tissues. The mRNA is quantitated precisely, e.g., fluorometrically. As a negative control, RNA was isolated from various normal tissues of the same tissue type as the cancerous tissues being tested.

5' nuclease assay data are initially expressed as Ct, or the threshold cycle. This is defined as the cycle at which the reporter signal accumulates above the background level of fluorescence. The ΔCt values are used as quantitative measurement of the relative number of starting copies of a particular target sequence in a nucleic acid sample when comparing cancer mRNA results to normal human mRNA results. As one Ct unit corresponds to 1 PCR cycle or approximately a 2-fold relative increase relative to normal, two units corresponds to a 4-fold relative increase, 3 units corresponds to an 8-fold relative increase and so on, one can quantitatively measure the relative fold increase in mRNA expression between two or more different tissues. Using this technique, the molecules listed below have been identified as being significantly overexpressed in a particular tumor(s) as compared to their normal non-cancerous counterpart tissue(s) (from both the same and different tissue donors) and thus, represent excellent polypeptide targets for the diagnosis and therapy of cancer in mammals.

Molecule	upregulation of expression in:	as compared to:
DNA77507 (TAT161)	breast tumor	normal breast tissue
DNA82343 (TAT157)	colon tumor	normal colon tissue
DNA88131 (TAT158)	breast tumor	normal breast tissue
DNA88131 (TAT158)	colon tumor	normal colon tissue
DNA95930 (TAT110)	colon tumor	normal colon tissue
DNA95930 (TAT110)	lung tumor	normal lung tissue
DNA95930 (TAT110)	prostate tumor	normal prostate tissue
DNA95930 (TAT110)	endometrial tumor	normal endometrial tissue
DNA95930 (TAT110)	ovarian tumor	normal ovarian tissue
DNA95930-1 (TAT210)	colon tumor	normal colon tissue
DNA95930-1 (TAT210)	lung tumor	normal lung tissue
DNA95930-1 (TAT210)	prostate tumor	normal prostate tissue
DNA95930-1 (TAT210)	endometrial tumor	normal endometrial tissue

Molecule	upregulation of expression in:	as compared to:
DNA95930-1 (TAT210)	ovarian tumor	normal ovarian tissue
DNA96930 (TAT112)	colon tumor	normal colon tissue
DNA96936 (TAT147)	colon tumor	normal colon tissue
DNA98591 (TAT162)	colon tumor	normal colon tissue
DNA108809 (TAT114)	kidney tumor	normal kidney tissue

DNA119488 (TAT119)	lung tumor	normal lung tissue
DNA188221 (TAT111)	colon tumor	normal colon tissue
DNA233876 (TAT146)	colon tumor	normal colon tissue
DNA193891 (TAT148)	colon tumor	normal colon tissue
DNA248170 (TAT187)	colon tumor	normal colon tissue
DNA194628 (TAT118)	kidney tumor	normal kidney tissue
DNA246415 (TAT167)	kidney tumor	normal kidney tissue
DNA210499 (TAT123)	lung tumor	normal lung tissue
DNA219894 (TAT211)	lung tumor	normal lung tissue
DNA215609 (TAT113)	colon tumor	normal colon tissue
DNA220432 (TAT128)	prostate tumor	normal prostate tissue
DNA226165 (TAT122)	lung tumor	normal lung tissue
DNA226237 (TAT117)	kidney tumor	normal kidney tissue
DNA246450 (TAT168)	kidney tumor	normal kidney tissue
DNA226456 (TAT144)	breast tumor	normal breast tissue
DNA237637 (TAT188)	breast tumor	normal breast tissue
DNA226539 (TAT126)	ovarian tumor	normal ovarian tissue
DNA236511 (TAT151)	ovarian tumor	normal ovarian tissue
DNA227224 (TAT121)	lung tumor	normal lung tissue
DNA247486 (TAT183)	lung tumor	normal lung tissue
DNA227800 (TAT131)	prostate tumor	normal prostate tissue
DNA228199 (TAT127)	ovarian tumor	normal ovarian tissue
DNA228199 (TAT127)	lung tumor	normal lung tissue
DNA228201 (TAT116)	colon tumor	normal colon tissue
DNA247488 (TAT189)	colon tumor	normal colon tissue
DNA236538 (TAT190)	colon tumor	normal colon tissue
DNA247489 (TAT191)	colon tumor	normal colon tissue
DNA228993 (TAT120)	lung tumor	normal lung tissue
DNA228994 (TAT124)	lung tumor	normal lung tissue
DNA236343 (TAT104)	breast tumor	normal breast tissue
DNA236534 (TAT102)	ovarian tumor	normal ovarian tissue
DNA246430 (TAT109)	breast tumor	normal breast tissue
DNA247480 (TAT142)	lung tumor	normal lung tissue
DNA98565 (TAT145)	glioma	normal brain tissue
DNA246435 (TAT152)	glioma	normal brain tissue
DNA226094 (TAT164)	glioma	normal brain tissue
DNA227578 (TAT165)	glioma	normal brain tissue

DNA231542 (TAT100)	glioma	normal brain tissue
DNA231542-1 (TAT284)	glioma	normal brain tissue
DNA231542-2 (TAT285)	glioma	normal brain tissue
DNA297393 (TAT285-1)	glioma	normal brain tissue

EXAMPLE 4: *In situ* Hybridization

In situ hybridization is a powerful and versatile technique for the detection and localization of nucleic acid sequences within cell or tissue preparations. It may be useful, for example, to identify sites of gene expression, analyze the tissue distribution of transcription, identify and localize viral infection, follow changes in specific mRNA synthesis and aid in chromosome mapping.

In situ hybridization was performed following an optimized version of the protocol by Lu and Gillett, Cell Vision 1:169-176 (1994), using PCR-generated ^{33}P-labeled riboprobes having homology to the target sequence to be detected. B riefly, formalin-fixed, paraffin-embedded human tissues were sectioned, deparaffinized, deproteinated in proteinase K (20 g/ml) for 15 minutes at 37°C, and further processed for *in situ* hybridization as described by Lu and Gillett, *supra*. A [^{33}-P] UTP-labeled antisense riboprobe was generated from a PCR product and hybridized at 55°C overnight. The slides were dipped in Kodak NTB2 nuclear track emulsion and exposed for 4 weeks.

^{33}P-Riboprobe synthesis

6.0 µl (125 mCi) of ^{33}P-UTP (Amersham BF 1002, SA<2000 Ci/mmol) were speed vac dried. To each tube containing dried ^{33}P-UTP, the following ingredients were added:

2.0 µl 5x transcription buffer

1.0 µl DTT (100 mM)

2.0 µl NTP mix (2.5 mM : 10 µ; each of 10 mM GTP, CTP & ATP + 10 µl H_2O)

1.0 µl UTP (50 µM)

1.0 µl Rnasin

1.0 µl DNA template (1µg)

1.0 µl H_2O

1.0 µl RNA polymerase (for PCR products T3 = AS, T7 = S, usually)

The tubes were incubated at 37°C for one hour. 1.0 µl RQ1 DNase were added, followed by incubation at 37°C for 15 minutes. 90 µl TE (10 mM Tris pH 7.6/1mM EDTA pH 8.0) were added, and the mixture was pipetted onto DE81 paper. The remaining solution was loaded in a Microcon-50 ultrafiltration unit, and spun using program 10 (6 minutes). The filtration unit was inverted over a second tube and spun using program 2 (3 minutes). After the final recovery spin, 100 µl TE were added. 1 µl of the final product was pipetted on DE81 paper and counted in 6 ml of Biofluor II.

The probe was run on a TBE/urea gel. 1-3 µl of the probe or 5 µl of RNA Mrk III were added to 3 µl of loading buffer. After heating on a 95°C heat block for three minutes, the probe was immediately placed on ice.

The wells of gel were flushed, the sample loaded, and run at 180-250 volts for 45 minutes. The gel was wrapped in saran wrap and exposed to XAR film with an intensifying screen in -70°C freezer one hour to overnight.

^{33}P-Hybridization

A. Pretreatment of frozen sections

The slides were removed from the freezer, placed on aluminium trays and thawed at room temperature for 5 minutes. The trays were placed in 55°C incubator for five minutes to reduce condensation. The slides were fixed for 10 minutes in 4% paraformaldehyde on ice in the fume hood, and washed in 0.5 x SSC for 5 minutes, at room temperature (25 ml 20 x SSC + 975 ml SQ H $_2$O). After deproteination in 0.5 µg/ml proteinase K for 10 minutes at 37°C (12.5 µl of 10 mg/ml stock in 250 ml prewarmed RNase-free RNAse buffer), the sections were washed in 0.5 x SSC for 10 minutes at room temperature. The sections were dehydrated in 70%, 95%, 100% ethanol, 2 minutes each.

B. Pretreatment of paraffin-embedded sections

The slides were deparaffinized, placed in SQ H_2O, and rinsed twice in 2 x SSC at room temperature, for 5 minutes each time. The sections were deproteinated in 20 µg/ml proteinase K (500 µl of 10 mg/ml in 250 ml RNase-free RNase buffer; 37°C, 15 minutes) - human embryo, or 8 x proteinase K (100 µl in 250 ml Rnase buffer, 37°C, 30 minutes) - formalin tissues. Subsequent rinsing in 0.5 x SSC and dehydration were performed as described above.

C. Prehybridization

The slides were laid out in a plastic box lined with Box buffer (4 x SSC, 50% formamide) - saturated filter paper.

D. Hybridization

1.0 x 10^6 cpm probe and 1.0 µl tRNA (50 mg/ml stock) per slide were heated at 95°C for 3 minutes. The slides were cooled on ice, and 48 µl hybridization buffer were added per slide. After vortexing, 50µl ^{33}P mix were added to 50 µl prehybridization on slide. The slides were incubated overnight at 55°C.

E. Washes

Washing was done 2 x 10 minutes with 2xSSC, EDTA at room temperature (400 ml 20 x SSC + 16 ml 0.25M EDTA, V_f=4L), followed by RNaseA treatment at 37°C for 30 minutes (500 µl of 10 mg/ml in 250 ml Rnase buffer = 20 µg/ml), The slides were washed 2 x 10 minutes with 2 x SSC, EDTA at room temperature. The stringency wash conditions were as follows: 2 hours at 55°C, 0.1 x SSC, EDTA (20 ml 20 x SSC + 16 ml EDTA, V_f=4L).

F. Oligonucleotides

In situ analysis was performed on a variety of DNA sequences disclosed herein. The oligonucleotides employed for these analyses were obtained so as to be complementary to the nucleic acids (or the complements thereof) as shown in the accompanying figures.

G. Results

In situ analysis was performed on a variety of DNA sequences disclosed herein. The results from these analyses are as follows.

(1) DNA95930 (TAT110)

In one analysis, significant expression is observed in 3/3 lung tumors, 3/3 colorectal adenocarcinomas, 1/1 prostate cancers, 3/3 transitional cell carcinomas and 3/3 endometrial adenocarcinomas, wherein the level of expression in the counterpart normal tissues is significantly less.

In a second independent analysis, significant expression is observed in 7/7 endometrial and 12/15 ovarian adenocarcinomas, wherein the level of expression in the counterpart normal tissues is significantly less.

In a third independent analysis, significant expression is observed in 24/26 colorectal tumor samples, wherein the level of expression in the counterpart normal tissue is significantly less.

Finally, in a fourth independent analysis, expression is observed in 8/26 samples of non-malignant prostate tissue, 55/82 samples of primary prostate cancer and in 5/23 samples of metastatic prostate cancer.

(2) DNA95930-1 (TAT210)

In one analysis, significant expression is observed in 3/3 lung tumors, 3/3 colorectal adenocarcinomas, 1/1 prostate cancers, 3/3 transitional cell carcinomas and 3/3 endometrial adenocarcinomas, wherein the level of expression in the counterpart normal tissues is significantly less.

In a second independent analysis, significant expression is observed in 7/7 endometrial and 12/15 ovarian adenocarcinomas, wherein the level of expression in the counterpart normal tissues is significantly less.

In a third independent analysis, significant expression is observed in 24/26 colorectal tumor samples, wherein the level of expression in the counterpart normal tissue is significantly less.

Finally, in a fourth independent analysis, expression is observed in 8/26 samples of non-malignant prostate tissue, 55/82 samples of primary prostate cancer and in 5/23 samples of metastatic prostate cancer.

(3) DNA96930 (TAT112)

Strong expression in colorectal cancers. Expression in the malignant epithelium appears significantly stronger than in adjacent benign epithelium. Additionally, strong expression is observed in all 23 of 23 samples of pancreatic adenocarcinoma tested, wherein expression in normal pancreatic tissue is not detectable.

(4) DNA96936 (TAT147)

In one analysis, a strongly positive signal was observed in 6/6 breast tumors. In another independent analysis, a positive signal was observed in 4/4 non small cell lung carcinomas, wherein the tumors appear to have stronger expression compared with normal lung. 1/1 endometrial adenocarcinomas shows strong expression and 3/3 colorectal adenocarcinomas show variable expression.

(5) DNA108809 (TAT114)

Positive signal in all renal cell carcinomas tested (n=3) while no expression observed in normal kidney tissue. Additionally, positive expression is observed in 5/12 stomach tumors, 5/24 colorectal tumors, 3/8 pancreatic tumors and 1/3 lung tumors. Normal non-cancerous tissue expression is limited to stomach and small intestine.

(6) DNA176766 (TAT132)

Positive signal in all endometrial adenocarcinomas tested (n=3) while no expression observed in normal endometrial tissue.

(7) DNA236463 (TAT150)

Positive signal in all endometrial adenocarcinomas tested (n=3) while no expression observed in normal endometrial tissue.

(8) DNA181162 (TAT129)

Neoplastic prostate epithelia are generally positive, with signal intensities varying from weak to strong between cases. Non-prostatic tissues are negative.

(9) DNA188221 (TAT111)

Strong signal seen in colonic multi-tumor array over malignant epithelium. In normal tissues, a certain probe gave specific signal over epithelial cells lining the lower 2/3 of the colonic crypts, the intensity of signal appeared significantly lower than in the colonic carcinomas. Positive expression is observed in 12/18 colorectal adenocarcinomas, 6/8 metastatic adenocarcinomas and 2/9 gastric adenocarcinomas.

(10) DNA233876 (TAT146)

Strong signal seen in colonic multi-tumor array over malignant epithelium. In normal tissues, a certain probe gave specific signal over epithelial cells lining the lower 2/3 of the colonic crypts, the intensity of signal appeared significantly lower than in the colonic carcinomas. Positive expression is observed in 12/18 colorectal adenocarcinomas, 6/8 metastatic adenocarcinomas and 2/9 gastric adenocarcinomas.

(11) DNA210499 (TAT123)

In one analysis, 12/14 ovarian adenocarcinomas are positive and 8/9 endometrial adenocarcinomas are positive. Normal ovarian stroma is negative as is uterine myometrium. Other normal ovarian and uterine tissues are negative.

In an independent analysis, 16/27 non small cell lung carcinomas are positive, wherein the signal is moderate or strong.

(12) DNA219894 (TAT211)

In one analysis, 12/14 ovarian adenocarcinomas are positive and 8/9 endometrial adenocarcinomas are positive. Normal ovarian stroma is negative as is uterine myometrium. Other normal ovarian and uterine tissues are negative.

In an independent analysis, 16/27 non small cell lung carcinomas are positive, wherein the signal is moderate or strong.

(13) DNA215609 (TAT113)

Strong signal seen in colonic carcinomas, with only very low level signal in normal colon. Lung and breast carcinomas were negative.

(14) DNA220432 (TAT128)

The only normal adult tissue expressing this gene is prostatic epithelium. The expression is of moderate to strong intensity and focal, it is more prevalent in hyperplastic epithelium.

In one analysis where 50 cases of primary prostate cancer are available for review, 29 cases (58%) are positive, 18 cases (36%) are negative and 3 cases (6%) are equivocal. In another analysis where 37 cases of primary prostate cancer are available for review, 33 cases (89%) are positive, 4 cases (11%) are negative. Finally, in another independent analysis where 27 cases of metastatic prostate cancer are available for review, 14 cases (52%) are positive, 11 cases (41%) are negative and 2 cases (7%) are equivocal.

(15) DNA226237 (TAT117)

In one analysis, two of 3 renal cell carcinomas are positive, wherein normal kidney expression is negative.

(16) DNA246450 (TAT168)

In one analysis, two of 3 renal cell carcinomas are positive, wherein normal kidney expression is negative.

(17) DNA227087 (TAT163)

A probe for this molecule showed a positive signal in a subpopulation of tumor-associated stromal cells in all tested cases of lung, breast, colon, pancreatic and endometrial carcinomas. The intensity of the labeling was often quite strong. In a case of colon adenocarcinoma with adjacent benign colon, labeling was restricted to the tumor-associated stroma and the normal benign tissue was negative. A breast fibroadenoma also showed labeling of subepithelial stromal cells.

(18) DNA266307 (TAT227)

A probe for this molecule showed a positive signal in a subpopulation of tumor-associated stromal cells in all tested cases of lung, breast, colon, pancreatic and endometrial carcinomas. The intensity of the labeling was often quite strong. In a case of colon adenocarcinoma with adjacent benign colon, labeling was restricted to the tumor-associated stroma and the normal benign tissue was negative. A breast fibroadenoma also showed labeling of subepithelial stromal cells.

(19) DNA266311 (TAT228)

A probe for this molecule showed a positive signal in a subpopulation of tumor-associated stromal cells in all tested cases of lung, breast, colon, pancreatic and endometrial carcinomas. The intensity of the labeling was often quite strong. In a case of colon adenocarcinoma with adjacent benign colon, labeling was restricted to the tumor-associated stroma and the normal benign tissue was negative. A breast fibroadenoma also showed labeling of subepithelial stromal cells.

(20) DNA266312 (TAT229)

A probe for this molecule showed a positive signal in a subpopulation of tumor-associated stromal cells in all tested cases of lung, breast, colon, pancreatic and endometrial carcinomas. The intensity of the labeling was often quite strong. In a case of colon adenocarcinoma with adjacent benign colon, labeling was restricted to the tumor-associated stroma and the normal benign tissue was negative. A breast fibroadenoma also showed labeling of subepithelial stromal cells.

(21) DNA266313 (TAT230)

A probe for this molecule showed a positive signal in a subpopulation of tumor-associated stromal cells in all tested cases of lung, breast, colon, pancreatic and endometrial carcinomas. The intensity of the labeling was often quite strong. In a case of colon adenocarcinoma with adjacent benign colon, labeling was restricted to the

tumor-associated stroma and the normal benign tissue was negative. A breast fibroadenoma also showed labeling of subepithelial stromal cells.

(22) DNA227224 (TAT121)

Expression is seen in 2 of 3 endometrial adenocarcinomas.

(23) DNA247486 (TAT183)

Expression is seen in 2 of 3 endometrial adenocarcinomas.

(24) DNA227800 (TAT131)

In one analysis, 46/64 primary prostate cancers are positive and 6/14 metastatic prostate cancers are positive. Weak to moderate expression is seen in prostate epithelium

(25) DNA228199 (TAT127)

Expression is observed in 13 of 15 ovarian tumors (adenocarcinoma and surface epithelial tumors). Benign ovarian surface epithelium is also positive. The expression level in most positive tumors is strong or moderate and fairly uniform. Expression is also observed in 8 of 9 uterine adenocarcinomas. Seven of 23 non small cell lung carcinomas are positive.

(26) DNA228201 (TAT116)

The malignant cells of 13/16 colorectal adenocarcinomas are positive for TAT116 expression. Additionally, 9/10 metastatic adenocarcinomas are positive for expression. Expression is also observed in the basal portions of normal colonic crypts.

(27) DNA247488 (TAT189)

The malignant cells of 13/16 colorectal adenocarcinomas are positive for TAT189 expression. Additionally, 9/10 metastatic adenocarcinomas are positive for expression. Expression is also observed in the basal portions of normal colonic crypts.

(28) DNA236538 (TAT190)

The malignant cells of 13/16 colorectal adenocarcinomas are positive for TAT190 expression. Additionally, 9/10 metastatic adenocarcinomas are positive for expression. Expression is also observed in the basal portions of normal colonic crypts.

(29) DNA247489 (TAT191)

The malignant cells of 13/16 colorectal adenocarcinomas are positive for TAT191 expression. Additionally, 9/10 metastatic adenocarcinomas are positive for expression. Expression is also observed in the basal portions of normal colonic crypts.

(30) DNA228994 (TAT124)

Thirteen of 61 cass of non small cell lung carcinoma are positive for expression of TAT124. Expression level in these positive tumor samples is significantly higher than in normal adult tissues.

(31) DNA231542 (TAT100)

In situ analysis performed as described above evidences significantly upregulated expression in human glioma and glioblastoma tissues as compared to normal brain (and other) tissue.

(32) DNA231542-1 (TAT284)

In situ analysis performed as described above evidences significantly upregulated expression in human glioma and gliobalstoma tissues as compared to normal brain (and other) tissue.

(33) DNA231542-2 (TAT285)

In situ analysis performed as described above evidences significantly upregulated expression in human glioma and glioblastoma tissues as compared to normal brain (and other) tissue.

(34) DNA297393 (TAT285-1)

In situ analysis performed as described above evidences significantly upregulated expression in human glioma and glioblastoma tissues as compared to normal brain (and other) tissue.

(35) DNA236534 (TAT102)

Expression of TAT102 is seen in 14 of 15 ovarian epithelial malignancies (adenocarcinoma, epithelial surface tumors, endometrioid Ca). Also, 8 of 9 endometrial adenocarcinomas of the uterus express TAT102. Moreover, expression of TAT102 is seen in 24 of 27 non-small cell lung cancers, positive cases include squamous and adenocarcinomas. Expression in these tumor tissues is significantly higher than in their normal tissue counterparts.

(36) DNA246430 (TAT109)

Fourteen of 92 breast tumor samples are positive for TAT109 expression. Expression in all normal tissues is undetectable.

(37) DNA264454 (TAT106)

Expression of TAT106 is observed in 38/88 breast tumors. Expression in normal breast tissue is weak or undetectable.

(38) DNA98565 (TAT145)

Positive signal for TAT145 was observed in most gliomas, glioblastomas, some melanomas, and normal brain (primarily localized to astrocytes). The signal intensity in the glioblastomas appeared to be greater than that in normal astrocytes. While the majority of glioma and glioblastoma samples tested were positive for TAT145 expression, the majority of normal brain samples tested were negative for such expression.

(39) DNA246435 (TAT152)

Positive signal for TAT152 was observed in most glioblastomas, some melanomas, and normal brain (primarily localized to astrocytes). The signal intensity in the glioblastomas appeared to be greater than that in normal astrocytes. While the majority of glioma and glioblastoma samples tested were positive for TAT152 expression, the majority of normal brain samples tested were negative for such expression.

(40) DNA167234 (TAT130)

Seventy cases of primary adenocarcinoma of the prostate were available for review. Of these 70 cases, 56 cases (80%) are positive for TAT130 expression. TAT130 expression in non-prostatic tissues is weak or undetectable.

(41) DNA235621 (TAT166)

Seventy cases of primary adenocarcinoma of the prostate were available for review. Of these 70 cases, 56 cases (80%) are positive for TAT166 expression. TAT166 expression in non-prostatic tissues is weak or undetectable.

(42) DNA236493 (TAT141)

Positive expression is observed in 70/148 breast carcinomas, 2/63 colorectal adenocarcinomas, 4/42 ovarian tumors, 9/69 non small cell lung carcinomas, 9/67 prostate adenocarcinomas and 5/25 gliomas. Expression in normal non-cancerous tissues appears restricted to prostate and breast epithelium.

(43) DNA226094 (TAT164)

Twenty one of 37 glioblastoma samples and 8 or 8 glioma samples were positive for TAT164 expression while all other tumor and normal tissues examined (including normal brain tissue) were negative.

(44) DNA227578 (TAT165)

Fifteen of 25 glioblastoma samples teste4d were positive for expression while significantly weaker expression was observed in the normal brain samples tested.

EXAMPLE 5: Immunohistochemistry Analysis

Antibodies against certain TAT polypeptides disclosed herein were prepared and immunohistochemistry analysis was performed as follows. Tissue sections were first fixed for 5 minutes in acetone/ethanol (frozen or paraffin-embedded). The sections were then washed in PBS and then blocked with avidin and biotin (Vector kit) for 10 minutes each followed by a wash in PBS. The sections were then blocked with 10% serum for 20 minutes and then blotted to remove the excess. A primary antibody was then added to the sections at a concentration of 10μg/ml for 1 hour and then the sections were washed in PBS. A biotinylated secondary antibody (anti-primary antibody) was then added to the sections for 30 minutes and then the sections were washed with PBS. The sections were then exposed to the reagents of the Vector ABC kit for 30 minutes and then the sections were washed in PBS. The sections were then exposed to Diaminobenzidine (Pierce) for 5 minutes and then washed in PBS. The sections were then counterstained with Mayers hematoxylin, covered with a coverslip and visualized. Immunohistochemistry analysis can also be performed as described in Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989 and Ausubel et al., Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). The results from these analyses are show below.

(1) DNA96930 (TAT112)

Significantly higher expression was detected in the apical surface of the colonic crypts of colon tumors than on the apical surface of the normal colonic crypts. Additionally, TAT112 was found to be significantly overexpressed in pancreatic adenocarcinoma cells as compared to normal pancreatic cells. Finally, IHC analysis performed as described above evidenced that TAT112 is significantly overexpressed in lung carcinoma as compared to normal lung tissue, non small cell lung carcinoma as compared to normal lung tissue and stomach carcinoma as compared to normal stomach tissue.

(2) DNA226539 (TAT126)

Positive expression is observed in 2/10 uterine adenocarcinomas, 9/17 ovarian adenocarcinomas and 2/20 non small cell lung carcinomas. Using this procedure, expression of TAT126 was not detectable in any normal tissue.

(3) DNA236511 (TAT151)

Positive expression is observed in 2/10 uterine adenocarcinomas, 9/17 ovarian adenocarcinomas and 2/20 non small cell lung carcinomas. Using this procedure, expression of TAT151 was not detectable in any normal tissue.

EXAMPLE 6: Verification and Analysis of Differential TAT Polypeptide Expression by GEPIS

TAT polypeptides which may have been identified as a tumor antigen as described in one or more of the above Examples were analyzed and verified as follows. An expressed sequence tag (EST) DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) was searched and interesting EST sequences were identified by GEPIS. Gene expression profiling*in silico* (GEPIS) is a bioinformatics tool developed at Genentech, Inc. that characterizes genes of interest for new cancer therapeutic targets. GEPIS takes advantage of large amounts of EST sequence and library information to determine gene expression profiles. GEPIS is capable of determining the expression profile of a gene based upon its proportional correlation with the number of its occurrences in EST databases, and it works by integrating the LIFESEQ® EST relational database and Genentech proprietary information in a stringent and statistically meaningful way. In this example, GEPIS is used to identify and cross-validate novel tumor antigens, although GEPIS can be configured to perform either very specific analyses or broad screening tasks. For the initial screen, GEPIS is used to identify EST sequences from the LIFESEQ® database that correlate to expression in a particular tissue or tissues of interest (often a tumor tissue of interest). The EST sequences identified in this initial screen (or consensus sequences obtained from aligning multiple related and overlapping EST sequences obtained from the initial screen) were then subjected to a screen intended to identify the presence of at least one transmembrane domain in the encoded protein. Finally, GEPIS was employed to generate a complete tissue expression profile for the various sequences of interest. Using this type of screening bioinformatics, various TAT polypeptides (and their encoding nucleic acid molecules) were identified as being significantly overexpressed in a particular type of cancer or certain cancers as compared to other cancers and/or normal non-cancerous tissues. The rating of GEPIS hits is based upon several criteria including, for example, tissue specificity, tumor specificity and expression level in normal essential and/or normal proliferating tissues. The following is a list of molecules whose tissue expression profile as determined by GEPIS evidences high tissue expression and significant upregulation of expression in a specific tumor or tumors as compared to other tumor(s) and/or normal tissues and optionally relatively low expression in normal essential and/or normal proliferating tissues. As such, the molecules listed below are excellent polypeptide targets for the diagnosis and therapy of cancer in mammals.

Molecule	upregulation of expression in:	as compared to:
DNA77507 (TAT161)	breast tumor	normal breast tissue
DNA77507 (TAT161)	colon tumor	normal colon tissue

DNA77507 (TAT161)	lung tumor	normal lung tissue
DNA77507 (TAT161)	kidney tumor	normal kidney tissue
DNA77507 (TAT161)	liver tumor	normal liver tissue
DNA77507 (TAT161)	ovarian tumor	normal ovarian tissue
DNA77507 (TAT161)	pancreatic tumor	normal pancreatic tissue
DNA77507 (TAT161)	rectum tumor	normal rectum tissue
DNA77507 (TAT161)	skin tumor	normal skin tissue
DNA77507 (TAT161)	uterine tumor	normal uterine tissue
DNA77507 (TAT161)	brain tumor	normal brain tissue
DNA77507 (TAT161)	soft tissue tumor	normal soft tissue
DNA77507 (TAT161)	bone tumor	normal bone tissue
DNA82343 (TAT157)	colon tumor	normal colon tissue
DNA82343 (TAT157)	ovarian tumor	normal ovarian tissue
DNA82343 (TAT157)	stomach tumor	normal stomach tissue
DNA82343 (TAT157)	thymus tumor	normal thymus tissue
DNA82343 (TAT157)	small intestine tumor	normal small intestine tissue
DNA87994 (TAT160)	breast tumor	normal breast tissue
DNA87994 (TAT160)	pancreatic tumor	normal pancreatic tissue
DNA87994 (TAT160)	colon tumor	normal colon tissue
DNA87994 (TAT160)	esophagus tumor	normal esophagus tissue
DNA87994 (TAT160)	ovarian tumor	normal ovarian tissue
DNA87994 (TAT160)	prostate tumor	normal prostate tissue
DNA88131 (TAT158)	breast tumor	normal breast tissue
DNA88131 (TAT158)	colon tumor	normal colon tissue
DNA88131 (TAT158)	lung tumor	normal lung tissue

Molecule	upregulation of expression in:	as compared to:
DNA88131 (TAT158)	pancreatic tumor	normal pancreatic tissue
DNA88131 (TAT158)	prostate tumor	normal prostate tissue
DNA88131 (TAT158)	stomach tumor	normal stomach tissue
DNA88131 (TAT158)	bladder tumor	normal bladder tissue
DNA88131 (TAT158)	brain tumor	normal brain tissue
DNA95930 (TAT110)	colon tumor	normal colon tissue
DNA95930 (TAT110)	lung tumor	normal lung tissue
DNA95930 (TAT110)	prostate tumor	normal prostate tissue
DNA95930 (TAT110)	endometrial tumor	normal endometrial tissue
DNA95930 (TAT110)	ovarian tumor	normal ovarian tissue

DNA95930 (TAT110)	breast tumor	normal breast tissue
DNA95930-1 (TAT210)	colon tumor	normal colon tissue
DNA95930-1 (TAT210)	lung tumor	normal lung tissue
DNA95930-1 (TAT210)	prostate tumor	normal prostate tissue
DNA95930-1 (TAT210)	endometrial tumor	normal endometrial tissue
DNA95930-1 (TAT210)	ovarian tumor	normal ovarian tissue
DNA95930-1 (TAT210)	breast tumor	normal breast tissue
DNA96917 (TAT159)	pancreatic tumor	normal pancreatic tissue
DNA96917 (TAT159)	lung tumor	normal lung tissue
DNA96917 (TAT159)	liver tumor	normal liver tissue
DNA96917 (TAT159)	prostate tumor	normal prostate tissue
DNA96930 (TAT112)	breast tumor	normal breast tissue
DNA96930 (TAT112)	colon tumor	normal colon tissue
DNA96930 (TAT112)	lung tumor	normal lung tissue
DNA96930 (TAT112)	ovarian tumor	normal ovarian tissue
DNA96930 (TAT112)	pancreatic tumor	normal pancreatic tissue
DNA96930 (TAT112)	stomach tumor	normal stomach tissue
DNA96936 (TAT147)	breast tumor	normal breast tissue
DNA96936 (TAT147)	colon tumor	normal colon tissue
DNA96936 (TAT147)	prostate tumor	normal prostate tissue
DNA96936 (TAT147)	uterine tumor	normal uterine tissue
DNA98565 (TAT145)	brain tumor	normal brain tissue
DNA98565 (TAT145)	colon tumor	normal colon tissue
DNA246435 (TAT152)	brain tumor	normal brain tissue
DNA246435 (TAT152)	colon tumor	normal colon tissue
DNA98591 (TAT162)	colon tumor	normal colon tissue
DNA98591 (TAT162)	small intestine tumor	normal small intestine tissue
DNA98591 (TAT162)	ovarian tumor	normal ovarian tissue
DNA98591 (TAT162)	esophagus tumor	normal esophagus tissue
DNA108809 (TAT114)	colon tumor	normal colon tissue
DNA108809 (TAT114)	lung tumor	normal lung tissue
DNA108809 (TAT114)	ovarian tumor	normal ovarian tissue
DNA108809 (TAT114)	brain tumor	normal brain tissue
DNA143493 (TAT103)	breast tumor	normal breast tissue
DNA167234 (TAT130)	prostate tumor	normal prostate tissue
DNA235621 (TAT166)	prostate tumor	normal prostate tissue
DNA176766 (TAT132)	kidney tumor	normal kidney tissue

DNA176766 (TAT132)	uterine tumor	normal uterine tissue
DNA236463 (TAT150)	kidney tumor	normal kidney tissue
DNA236463 (TAT150)	uterine tumor	normal uterine tissue
DNA181162 (TAT129)	prostate tumor	normal prostate tissue
DNA188221 (TAT111)	colon tumor	normal colon tissue
DNA188221 (TAT111)	liver tumor	normal liver tissue
DNA188221 (TAT111)	lung tumor	normal lung tissue
DNA233876 (TAT146)	colon tumor	normal colon tissue

Molecule	upregulation of expression in:	as compared to:
DNA233876 (TAT146)	liver tumor	normal liver tissue
DNA233876 (TAT146)	lung tumor	normal lung tissue
DNA193891 (TAT148)	prostate tumor	normal prostate tissue
DNA193891 (TAT148)	breast tumor	normal breast tissue
DNA248170 (TAT187)	breast tumor	normal breast tissue
DNA248170 (TAT187)	prostate tumor	normal prostate tissue
DNA194628 (TAT118)	kidney tumor	normal kidney tissue
DNA246415 (TAT167)	kidney tumor	normal kidney tissue
DNA215609 (TAT113)	colon tumor	normal colon tissue
DNA220432 (TAT128)	prostate tumor	normal prostate tissue
DNA226094 (TAT164)	breast tumor	normal breast tissue
DNA226094 (TAT164)	brain tumor	normal brain tissue
DNA226094 (TAT164)	ovarian tumor	normal ovarian tissue
DNA226094 (TAT164)	lung tumor	normal lung tissue
DNA226165 (TAT122)	breast tumor	normal breast tissue
DNA226165 (TAT122)	endometrial tumor	normal endometrial tissue
DNA226165 (TAT122)	lung tumor	normal lung tissue
DNA226165 (TAT122)	colon tumor	normal colon tissue
DNA226237 (TAT117)	kidney tumor	normal kidney tissue
DNA246450 (TAT168)	kidney tumor	normal kidney tissue
DNA246450 (TAT168)	brain tumor	normal brain tissue
DNA226456 (TAT144)	breast tumor	normal breast tissue
DNA226456 (TAT144)	brain tumor	normal brain tissue
DNA226456 (TAT144)	endometrial tumor	normal endometrial tissue
DNA226456 (TAT144)	kidney tumor	normal kidney tissue
DNA226456 (TAT144)	lung tumor	normal lung tissue
DNA237637 (TAT188)	breast tumor	normal breast tissue

DNA237637 (TAT188)	brain tumor	normal brain tissue
DNA237637 (TAT188)	endometrial tumor	normal endometrial tissue
DNA237637 (TAT188)	kidney tumor	normal kidney tissue
DNA237637 (TAT188)	lung tumor	normal lung tissue
DNA226539 (TAT126)	colon tumor	normal colon tissue
DNA226539 (TAT126)	endometrial tumor	normal endometrial tissue
DNA226539 (TAT126)	ovarian tumor	normal ovarian tissue
DNA226539 (TAT126)	pancreatic tumor	normal pancreatic tissue
DNA236511 (TAT151)	colon tumor	normal colon tissue
DNA236511 (TAT151)	endometrial tumor	normal endometrial tissue
DNA236511 (TAT151)	ovarian tumor	normal ovarian tissue
DNA236511 (TAT151)	pancreatic tumor	normal pancreatic tissue
DNA226771 (TAT115)	colon tumor	normal colon tissue
DNA227087 (TAT163)	breast tumor	normal breast tissue
DNA227087 (TAT163)	colon tumor	normal colon tissue
DNA227087 (TAT163)	endocrine tumor	normal endocrine tissue
DNA227087 (TAT163)	kidney tumor	normal kidney tissue
DNA227087 (TAT163)	liver tumor	normal liver tissue
DNA227087 (TAT163)	lung tumor	normal lung tissue
DNA227087 (TAT163)	pancreatic tumor	normal pancreatic tissue
DNA227087 (TAT163)	uterine tumor	normal uterine tissue
DNA227087 (TAT163)	prostate tumor	normal prostate tissue
DNA227087 (TAT163)	bladder tumor	normal bladder tissue
DNA266307 (TAT227)	breast tumor	normal breast tissue
DNA266307 (TAT227)	colon tumor	normal colon tissue
DNA266307 (TAT227)	endocrine tumor	normal endocrine tissue
DNA266307 (TAT227)	kidney tumor	normal kidney tissue
DNA266307 (TAT227)	liver tumor	normal liver tissue

Molecule	upregulation of expression in:	as compared to:
DNA266307 (TAT227)	lung tumor	normal lung tissue
DNA266307 (TAT227)	pancreatic tumor	normal pancreatic tissue
DNA266307 (TAT227)	uterine tumor	normal uterine tissue
DNA266307 (TAT227)	prostate tumor	normal prostate tissue
DNA266307 (TAT227)	bladder tumor	normal bladder tissue
DNA266311 (TAT228)	breast tumor	normal breast tissue
DNA266311 (TAT228)	colon tumor	normal colon tissue

DNA266311 (TAT228)	endocrine tumor	normal endocrine tissue
DNA266311 (TAT228)	kidney tumor	normal kidney tissue
DNA266311 (TAT228)	liver tumor	normal liver tissue
DNA266311 (TAT228)	lung tumor	normal lung tissue
DNA266311 (TAT228)	pancreatic tumor	normal pancreatic tissue
DNA266311 (TAT228)	uterine tumor	normal uterine tissue
DNA266311 (TAT228)	prostate tumor	normal prostate tissue
DNA266311 (TAT228)	bladder tumor	normal bladder tissue
DNA266312 (TAT229)	breast tumor	normal breast tissue
DNA266312 (TAT229)	colon tumor	normal colon tissue
DNA266312 (TAT229)	endocrine tumor	normal endocrine tissue
DNA266312 (TAT229)	kidney tumor	normal kidney tissue
DNA266312 (TAT229)	liver tumor	normal liver tissue
DNA266312 (TAT229)	lung tumor	normal lung tissue
DNA266312 (TAT229)	pancreatic tumor	normal pancreatic tissue
DNA266312 (TAT229)	uterine tumor	normal uterine tissue
DNA266312 (TAT229)	prostate tumor	normal prostate tissue
DNA266312 (TAT229)	bladder tumor	normal bladder tissue
DNA266313 (TAT230)	breast tumor	normal breast tissue
DNA266313 (TAT230)	colon tumor	normal colon tissue
DNA266313 (TAT230)	endocrine tumor	normal endocrine tissue
DNA266313 (TAT230)	kidney tumor	normal kidney tissue
DNA266313 (TAT230)	liver tumor	normal liver tissue
DNA266313 (TAT230)	lung tumor	normal lung tissue
DNA266313 (TAT230)	pancreatic tumor	normal pancreatic tissue
DNA266313 (TAT230)	uterine tumor	normal uterine tissue
DNA266313 (TAT230)	prostate tumor	normal prostate tissue
DNA266313 (TAT230)	bladder tumor	normal bladder tissue
DNA227224 (TAT121)	breast tumor	normal breast tissue
DNA227224 (TAT121)	endometrial tumor	normal endometrial tissue
DNA227224 (TAT121)	lung tumor	normal lung tissue
DNA227224 (TAT121)	skin tumor	normal skin tissue
DNA247486 (TAT183)	breast tumor	normal breast tissue
DNA247486 (TAT183)	endometrial tumor	normal endometrial tissue
DNA247486 (TAT183)	lung tumor	normal lung tissue
DNA247486 (TAT183)	skin tumor	normal skin tissue
DNA227578 (TAT165)	brain tumor	normal brain tissue

DNA227800 (TAT131)	prostate tumor	normal prostate tissue
DNA227800 (TAT131)	kidney tumor	normal kidney tissue
DNA227904 (TAT140)	breast tumor	normal breast tissue
DNA228199 (TAT127)	uterine tumor	normal uterine tissue
DNA228199 (TAT127)	fallopian tube tumor	normal fallopian tube tissue
DNA228199 (TAT127)	ovarian tumor	normal ovarian tissue
DNA228199 (TAT127)	lung tumor	normal lung tissue
DNA228201 (TAT116)	colon tumor	normal colon tissue
DNA247488 (TAT189)	colon tumor	normal colon tissue
DNA236538 (TAT190)	colon tumor	normal colon tissue
DNA247489 (TAT191)	colon tumor	normal colon tissue

Molecule	upregulation of expression in:	as compared to:
DNA231312 (TAT143)	colon tumor	normal colon tissue
DNA231542 (TAT100)	brain tumor	normal brain tissue
DNA231542 (TAT100)	glioma	normal glial tissue
DNA231542-1 (TAT284)	brain tumor	normal brain tissue
DNA231542-1 (TAT284)	glioma	normal glial tissue
DNA231542-2 (TAT285)	brain tumor	normal brain tissue
DNA231542-2 (TAT285)	glioma	normal glial tissue
DNA297393 (TAT285-1)	brain tumor	normal brain tissue
DNA297393 (TAT285-1)	glioma	normal glial tissue
DNA232754 (TAT125)	lung tumor	normal lung tissue
DNA236246 (TAT153)	breast tumor	normal breast tissue
DNA236343 (TAT104)	breast tumor	normal breast tissue
DNA236493 (TAT141)	breast tumor	normal breast tissue
DNA236493 (TAT141)	glioblastoma tumor	normal glial tissue
DNA236534 (TAT102)	breast tumor	normal breast tissue
DNA236534 (TAT102)	lung tumor	normal lung tissue
DNA236534 (TAT102)	pancreatic tumor	normal pancreatic tissue
DNA236534 (TAT102)	prostate tumor	normal prostate tissue
DNA236534 (TAT102)	bladder tumor	normal bladder tissue
DNA247480 (TAT142)	lung tumor	normal lung tissue
DNA264454 (TAT106)	breast tumor	normal breast tissue
DNA264454 (TAT106)	prostate tumor	normal prostate tissue
DNA264454 (TAT106)	ovarian tumor	normal ovarian tissue

EXAMPLE 7: Use of TAT as a hybridization probe

The following method describes use of a nucleotide sequence encoding TAT as a hybridization probe for, i.e., diagnosis of the presence of a tumor in a mammal.

DNA comprising the coding sequence of full-length or mature TAT as disclosed herein can also be employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of TAT) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled TAT-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42 °C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence TAT can then be identified using standard techniques known in the art.

EXAMPLE 8: Expression of TAT in *E. coli*

This example illustrates preparation of an unglycosylated form of TAT by recombinant expression in *E. coli*.

The DNA sequence encoding TAT is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the TAT coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., supra. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized TAT protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

TAT may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding TAT is initially amplified using selected PCR primers. The primers will contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an expression vector, which is used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D.600 of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g $(NH_4)_2SO_4$, 0.71 g sodium citrate•2H2O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM $MgSO_4$) and grown for approximately 20-30 hours at 30°C with shaking. Samples are removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

E. coli paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged at 40,000 rpm in a Beckman Ultracentifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The clarified extract is loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein are pooled and stored at 4°C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The refolding solution is stirred gently at 4°C for 12-36 hours. The refolding reaction is quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution is filtered through a 0.22 micron filter and acetonitrile is added to 2-10% final concentration. The refolded protein is chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance are analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein are pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

Fractions containing the desired folded TAT polypeptide are pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins are formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

Certain of the TAT polypeptides disclosed herein have been successfully expressed and purified using this technique(s).

EXAMPLE 9: Expression of TAT in mammalian cells

This example illustrates preparation of a potentially glycosylated form of TAT by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the TAT DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the TAT DNA using ligation methods such as described in Sambrook et al., supra. The resulting vector is called pRK5-TAT.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10μg pRK5-TAT DNA is mixed with about 1 μg DNA encoding the VA RNA gene [Thimmappaya et al., Cell, 31:543 (1982)] and dissolved in 500 μl of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M $CaCl_2$. To this mixture is added, dropwise, 500μl of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM $NaPO_4$, and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 μCi/ml ^{35}S-cysteine and 200 μCi/ml ^{35}S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of TAT polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, TAT may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac et al., Proc. Natl. Acad. Sci., 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μg pRK5-TAT DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μg/ml bovine insulin and 0.1 μg/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris.

The sample containing expressed TAT can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, TAT can be expressed in CHO cells. The pRK5-TAT can be transfected into CHO cells using known reagents such as $CaPO_4$ or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ^{35}S-methionine. After determining the presence of TAT polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed TAT can then be concentrated and purified by any selected method.

Epitope-tagged TAT may also be expressed in host CHO cells. The TAT may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged TAT insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged TAT can then be concentrated and purified by any selected method, such as by Ni^{2+}-chelate affinity chromatography.

TAT may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins are fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas et al., Nucl. Acids Res. 24:9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

Twelve micrograms of the desired plasmid DNA is introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect® (Quiagen), Dosper® or Fugene® (Boehringer Mannheim). The cells are grown as described in Lucas et al. supra. Approximately 3 x 10^7 cells are frozen in an ampule for further growth and production as described below.

The ampules containing the plasmid DNA are thawed by placement into water bath and mixed by vortexing. The contents are pipetted into a centrifuge tube containing 10 mLs of media and centrifuged at 1000 rpm for 5 minutes. The supernatant is aspirated and the cells are resuspended in 10 mL of selective media (0.2 μm filtered PS20 with 5% 0.2 μm diafiltered fetal bovine serum). The cells are then aliquoted into a 100 mL spinner

containing 90 mL of selective media. After 1-2 days, the cells are transferred into a 250 mL spinner filled with 150 mL selective growth medium and incubated at 37°C. After another 2-3 days, 250 mL, 500 mL and 2000 mL spinners are seeded with 3 x 10^5 cells/mL. The cell media is exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in U.S. Patent No. 5,122,469, issued June 16, 1992 may actually be used. A 3L production spinner is seeded at 1.2 x 10^6 cells/mL. On day 0, the cell number pH ie determined. On day 1, the spinner is sampled and sparging with filtered air is commenced. On day 2, the spinner is sampled, the temperature shifted to 33°C, and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion) taken. Throughout the production, the pH is adjusted as necessary to keep it at around 7.2. After 10 days, or until the viability dropped below 70%, the cell culture is harvested by centrifugation and filtering through a 0.22 μm filter. The filtrate was either stored at 4°C or immediately loaded onto columns for purification.

For the poly-His tagged constructs, the proteins are purified using a Ni-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4 °C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc-containing) constructs are purified from the conditioned media as follows. The conditioned medium is pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275 μL of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

Certain of the TAT polypeptides disclosed herein have been successfully expressed and purified using this technique(s).

EXAMPLE 10: Expression of TAT in Yeast

The following method describes recombinant expression of TAT in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of TAT from the ADH2/GAPDH promoter. DNA encoding TAT and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of TAT. For secretion, DNA encoding TAT can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native TAT signal peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of TAT.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant TAT can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing TAT may further be purified using selected column chromatography resins.

Certain of the TAT polypeptides disclosed herein have been successfully expressed and purified using this technique(s).

EXAMPLE 11: Expression of TAT in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of TAT in Baculovirus-infected insect cells.

The sequence coding for TAT is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding TAT or the desired portion of the coding sequence of TAT such as the sequence encoding an extracellular domain of a transmembrane protein or the sequence encoding the mature protein if the protein is extracellular is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression are performed as described by O'Reilley et al., Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-his tagged TAT can then be purified, for example, by $^{2+}$Ni chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al Nature, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM $MgCl_2$; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45 μm filter. A Ni^{2+}-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A_{280} with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A_{280} baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by

SDS-PAGE and silver staining or Western blot with Ni^{2+}-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His_{10}-tagged TAT are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) TAT can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

Certain of the TAT polypeptides disclosed herein have been successfully expressed and purified using this technique(s).

EXAMPLE 12: Preparation of Antibodies that Bind TAT

This example illustrates preparation of monoclonal antibodies which can specifically bind TAT. This example further illustrates preparation of monoclonal antibodies which can specifically bind TAT188(E16) polypeptide.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, supra. Immunogens that may be employed include purified TAT, fusion proteins containing TAT, and cells expressing recombinant TAT on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the TAT immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-TAT antibodies.

The TAT188 polypeptide, E16, is a twelve-transmembrane protein with both the C- and N-terminus located intracellularly. TAT(E16) is a subunit of the Na2+ independent large neutral amino acid transporter, heterodimerized with a common heavy chain 4F2hc (CD98hc) to form a functional unit (see diagram in Fig. 155). To prepare anti-TAT188 (anti-E16) antibodies targeting extracellular domain(s), PC3 cells, which endogenously express E16, were used as the immunogen. Briefly, PC3 cells (22 x 10^6 cells/ml) were injected into Balb-c mice. Antibodies titers were tested against control 293 cells and 293-E16 cells (293 cells transfected with and producing E16). Six hybridomas were prepared which expressed anti-E16 antibodies. Of these, three antibodies were shown to have good specificity for E16 by FACS analysis (see Fig. 156) in which cell surface binding paralleled E16 surface expression. Specifically E16 siRNA transfection inhibited E16 expression (see Example 17 herein) and also decreased the amount of anti-E16 antibody bound to the surface of PC3 cells. These antibodies are referred to herein as 3B5.1 (or 3B5), 12G12.1 (or 12G12), and 12B9.1 (or 12B9) and were used for further experiments disclosed herein.

Hybridomas expressing the antibodies of the invention were prepared as follows. After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of TAT. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen

cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells were screened in an ELISA for reactivity against TAT. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against TAT is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-TAT monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

Antibodies directed against certain of the TAT polypeptides disclosed herein have been successfully produced using this technique(s). More specifically, functional monoclonal antibodies that are capable of recognizing and binding to TAT protein (as measured by standard ELISA, FACS sorting analysis and/or immunohistochemistry analysis) have been successfully generated against the following TAT proteins as disclosed herein: TAT110 (DNA95930), TAT210 (DNA95930-1), TAT113 (DNA215609), TAT126 (DNA226539), TAT151 (DNA236511), TAT111 (DNA188221), TAT146 (DNA233876), TAT112 (DNA96930), TAT145 (DNA98565), TAT152 (DNA246435), TAT141 (DNA236493), TAT114 (DNA108809), TAT104 (DNA236343), TAT100 (DNA231542), TAT284 (DNA231542-1), TAT285 (DNA231542-2), TAT285-1 (DNA297393), TAT144 (DNA226456), TAT188 (DNA237637), TAT123 (DNA210499), TAT211 (DNA219894), TAT102 (DNA236534), TAT127 (DNA228199) and TAT128 (DNA220432). Interestingly, Applicants have identified that the monoclonal antibodies prepared against TAT111 (DNA188221) and TAT146 (DNA233876) are capable of blocking activation of the EphB2R receptor encoded by the DNA188221 and DNA233876 molecules by its associated ligand polypeptide. As such, antibodies and methods for using those antibodies for blocking activation of the EphB2R receptor (i.e., TAT111 and TAT146 polypeptides) by its associated ligand are encompassed within the presently described invention. Moreover, Applicants have identified that monoclonal antibodies directed against the TAT110 (DNA95930) and TAT210 (DNA95930-1) polypeptides (i.e., IL-20 receptor alpha polypeptides) are capable of inhibiting activation of the IL20 receptor alpha by IL-19 protein.As such, antibodies and methods for using those antibodies for inhibiting activation of the IL-20 receptor alpha (i.e., TAT110 and TAT210 polypeptides) by IL-19 are encompassed within the presently described invention.

In addition to the successful preparation of monoclonal antibodies directed against the TAT polypeptides as described herein, many of those monoclonal antibodies have been successfully conjugated to a cell toxin for use in directing the cellular toxin to a cell (or tissue) that expresses a TAT polypeptide of interested (both *in vitro* and *in vivo*). For example, toxin (e.g., DM1) derivitized monoclonal antibodies have been successfully generated to the following TAT polypeptides as described herein: TAT110 (DNA95930), TAT210 (DNA95930-1), TAT112 (DNA96930), TAT113 (DNA215609), TAT111 (DNA188221) and TAT146 (DNA233876). As shown herein below, other useful conjugate toxins include the auristatins MMAE and MMAF.

EXAMPLE 13: Purification of TAT Polypeptides Using Specific Antibodies

Native or recombinant TAT polypeptides may be purified by a variety of standard techniques in the art of protein purification. For example, pro-TAT polypeptide, mature TAT polypeptide, or pre-TAT polypeptide is purified by immunoaffinity chromatography using antibodies specific for the TAT polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-TAT polypeptide antibody to an activated chromatographic resin.

Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such an immunoaffinity column is utilized in the purification of TAT polypeptide by preparing a fraction from cells containing TAT polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble TAT polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

A soluble TAT polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of TAT polypeptide (*e.g.*, high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/TAT polypeptide binding (*e.g.*, a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and TAT polypeptide is collected.

EXAMPLE 14: *In Vitro* Tumor Cell Killing Assay

Mammalian cells expressing the TAT polypeptide of interest may be obtained using standard expression vector and cloning techniques. Alternatively, many tumor cell lines expressing TAT polypeptides of interest are publicly available, for example, through the ATCC and can be routinely identified using standard ELISA or FACS analysis. Anti-TAT polypeptide monoclonal antibodies (and toxin conjugated derivatives thereof) may then be employed in assays to determine the ability of the antibody to kill TAT polypeptide expressing cells *in vitro*.

For example, cells expressing the TAT polypeptide of interest are obtained as described above and plated into 96 well dishes. In one analysis, the antibody/toxin conjugate (or naked antibody) is included throughout the cell incubation for a period of 4 days. In a second independent analysis, the cells are incubated for 1 hour with the antibody/toxin conjugate (or naked antibody) and then washed and incubated in the absence of antibody/toxin conjugate for a period of 4 days. Cell viability is then measured using the CellTiter-Glo Luminescent Cell Viability Assay from Promega (Cat# G7571). Untreated cells serve as a negative control.

In one specific analysis, the ability of monoclonal antibodies directed against TAT112 (DNA96930) were analyzed for the ability to kill cells expressing that polypeptide. In one analysis, an expression vector called gD.NCA was prepared. The TAT112 polypeptide encoding sequences inserted into that vector are driven by an SV40 promoter and the vector also contains the SV40 early poly A signal. The gD.NCA vector was co-transfected into PC3 cells along with an SV40 vector that expresses Neo resistance in PC3 cells, and positive transformants were selected in 800 μg/ml G418. Positive clones were isolated in 96 well plates and analyzed by flow cytometry using an anti-TAT112 monoclonal antibody prepared as described above and called 3E6. Clone 3 was selected for the analysis as it was found to express a high level of TAT112 polypeptide on its surface. In a second independent analysis, the pancreatic cancer cell line, Hpaf II, was obtained from the ATCC and employed in the assay.

In another specific analysis, the ability of monoclonal antibodies directed against TAT188 (DNA237637) were analyzed for the ability to kill cells expressing that polypeptide.

First, the ability of a naked anti-TAT188(E16) antibody to affect cellular activity was demonstrated.

A. Anti-TAT188(E16) monoclonal antibodies stimulate TAT188(E16) internalization and reduce amino acid transport activity.

The binding of anti-E16 antibody to cells expressing the E16 protein elicited antibody internalization which paralleled a reduction in amino acid transport. PC3 cells were incubated with primary antibodies for indicated time with proteasome inhibitors (Sigma) at 37°C. Then antibodies are removed and cells are washed with PBS several times. Cells are fixed with 4% PFA, then permeabilized with PBS/0.1%TritonX-100. After blocking with 10% goat serum, cells were stained with secondary antibody (goat anti-mouse IgG-Cy3 conjugated (Jackson immunolabs)) and analyzed by fluorescent microscope. The results show that an anti-E16 monoclonal antibody binds and is internalized by PC3 cells expressing TAT188(E16) on their surface (see Fig. 157). The demonstration of anti-TAT188 antibody internalization highlights a useful feature of the antibody-E16 interaction because the internalization of a bound antibody can result in the uptake of toxins conjugated to the anti-TAT188 antibodies (see sections B and C of this Example 14).

The function of TAT188(E16) as an amino acid transport protein was inhibited by the binding of anti-TAT188(E16) antibody. The leucine transport functional assay was performed as follows. Cells were grown in 24 well plate and treated with anti-E16 antibody for over night. Cells were then rinsed with warm Na++ free Hepes-Ringer, and incubated for 10min at 37°C. Buffer was changed to 200 μl/well of assay buffer containing radiolabeled amino acid (L-[2, 3-^{3}H]-Arginine or L-[3, 4, 5-^{3}H(N)]-Leucine) 5 μCi/ml, 50 μl insodium-free Hepes-Ringer). After incubating the radioactive mixture for30 sec at 37°C, buffer was removed by aspiration and cells were wash with ice-cold sodium-free Hepes-Ringer 1ml/well for 4 times. Thep lates were dried and to them was added 0.2ml/well of 0.2%SDS/0.2N NaOH. 0.1ml aliquots of lysates from each sample were neutralized with 0.1ml of 0.2N Hcl, after which liquid scintillation cocktail was added. ([^{3}H]) radioactivity of the lysates was tested using scintillation counter. The results in Figure 158 also demonstrate that anti-E16 antibody binding to E16 protein on the cell surface causes overall reduction in amino acid transport after approximately 24 hours. The reduction is due to concomitant E16 polypeptide internalization and loss of

available E16 for transport function.

B. The binding of toxin-conjugated anti-TAT188(E16) antibodies to TAT188(E16)-expressing cells results in cell killing.

Preparation of anti-TAT188-MC-vc-PAB-MMAE by conjugation of anti-TAT188 and MC-vc-PAB-MMAE

Antibodies provide a convenient and highly specific method for delivering cytotoxins to cells where a relatively unique feature of the cell, such as a tissue specific (or disease specific or both) surface protein, allows the cell to be targeted for killing while avoiding the killing of healthy cells. The auristatin cytotoxins coupled to an antibody via a MC-vc-PAB linker as disclosed hereinabove, are useful to demonstrate cell killing resulting from E16 mediated internalization of the antibody-toxin conjugate. The procedure was performed as follows. Using a MC-vc-PAB linker, the MMAE toxin was covalently linked to cysteine residues on anti-TAT188 antibodies. Anti-TAT188, dissolved in 500mM sodium borate and 500 mM sodium chloride at pH 8.0 is treated with an excess of 100mM dithiothreitol (DTT). After incubation at 37 °C for about 30 minutes, the buffer is exchanged by elution over Sephadex G25 resin and eluted with PBS with 1mM DTPA. The thiol/Ab value is checked by determining the reduced antibody concentration from the absorbance at 280 nm of the solution and the thiol concentration by reaction with DTNB (Aldrich, Milwaukee, WI) and determination of the absorbance at 412 nm. The reduced antibody dissolved in PBS is chilled on ice.

The drug linker reagent, maleimidocaproyl-valine-citrulline-PAB-monomethyl auristatin E (MMAE), i.e. MC-vc-PAB-MMAE, dissolved in DMSO, is diluted in acetonitrile and water at known concentration, and added to the chilled reduced anti-TAT188(E16) antibody in PBS. After about one hour, an excess of maleimide is added to quench the reaction and cap any unreacted antibody thiol groups. The reaction mixture is concentrated by centrifugal ultrafiltration and anti-TAT188(E16)-MC-vc-PAB-MMAE was purified and desalted by elution through G25 resin in PBS, filtered through 0.2 μm filters under sterile conditions, and frozen for storage.

Preparation of anti-TAT188-MC-vc-PAB-MMAF by conjugation of anti-TAT188 and MC-vc-PAB-MMAF

Anti-TAT188-MC-val-cit-PAB-MMAF was prepared by conjugation of anti-TAT188(E16) and MC-val-cit-PAB-MMAF following the procedure disclosed above for conjugation of MMAE. MMAF is a derivative of MMAE with a phenylalanine at the C-terminus of the drug.

Determination of in vitro Cytotoxicity of toxin-conjugated anti-TAT188

The toxin-conjugated 3B5, 12B9, and 12G12 antibodies for this experiment were prepared as 3B5-MC-vc-PAB-MMAE, 12B9-MC-vc-PAB-MMAE, and 3B5-MC-vc-PAB-MMAE conjugates, respectively. Antibody-toxin conjugate, alpha-IL8-MC-vc-PAB-MMAE, was used as a negative control. PC3 cells (a prostate cancer cell line, available from ATCC and which endogenously expressed E16 on its surface) and

Colo205 cells (a human colon cancer cell line, available from ATCC and which endogenously expresses E16 on its surface) were contacted with increasing concentrations of toxin-conjugated antibodies and monitored for cell killing by a standard cell viability bioluminescence assay (CellTiterGlo™ Kit, Promega) as described above. The results shown in Fig. 159A (PC3 cells) and Fig. 159B (Colo205 cells) demonstrate that the three toxin-conjugated anti-E16 antibodies promote killing of cells that express E16.

Another toxin, MMAF, was linked to antibody and tested for cytotoxicity. The toxin-linked 3B5 antibodies for this experiment were designated 3B5-MC-vc-PAB-MMAE and 3B5-MC-vc-PAB-MMAF. Antibody-toxin conjugate, alpha-IL8-vc-MMAE, was used as a negative control. The average number of toxin molecules attached to each antibody was approximately four toxin molecules per antibody as shown in Fig. 160. Human colon cancer cell line, Colo205, cells were seeded in 96well plates at 4000cells/well in 50 µl of culture medium. The next day cells were incubated with 3B5-MC-vc-PAB-MMAE, 3B5-MC-vc-PAB-MMAF or alpha-IL8-vc-MMAE antibodies in graded concentrations diluted in 50 µl of culture medium for 72hrs. Cell viability was measured using a CellTiter-Glo™ luminescent cell viability assay kit (Promega). The results shown in Fig. 160 demonstrate that MC-vc-PAB-MMAE and MC-vc-PAB-MMAF-conjugated anti-E16 antibodies have the ability to kill cells expressing E16, with MMAF toxin (EC50 approximately 0.06 µg/ml) providing greater cell killing than MMAF (EC50 approximately 0.007 µg/ml).

Development of the antibodies of the invention as therapeutic agent in humans requires early toxicity studies in monkeys and/or primates. The E16 amino acid sequence varies across species such as human versus monkey, rat and mouse E16 amino acid sequences is 96.26%, 91.05% and 90.66, respectively. To determine whether the anti-human TAT188(E16) antibodies were also capable of binding and killing monkey cells, the following in vitro experiment was performed. A monkey cell line, COS7 (available from ATCC), that endogenously expresses monkey E16 on its surface was used. The anti-E16 antibodies 3B5, 12B9, and 12G12 were contacted with human, monkey, and mouse cells endogenously expressing human, monkey and mouse E16, respectively, and analyzed by standard FACS analysis. The FACS showed that each of the three anti-human E16 antibodies 3B5, 12G12, and 12B9 bind monkey COS7 cells expressing human E16 (see Fig. 161A) as well as human cells expressing human E16 (see Fig. 161B), but do not bind mouse E16 expressing cells (see Fig. 161B). Thus, the anti-TAT188(E16) antibodies of the invention are useful for clinical activity, efficacy and toxicity studies.

EXAMPLE 15: *In Vivo* Tumor Cell Killing Assay

Anti-TAT112 activity:

To test the efficacy of unconjugated anti-TAT112 monoclonal antibodies, anti-TAT112 antibody was injected intraperitoneally into nude mice 24 hours prior to receiving PC3.gD.NCA clone 3 cells (obtained as described in Example 14 above) subcutaneously in the flank. Antibody injections continued twice per week for the remainder of the study. Tumor volume was measured twice per week.

To test the efficacy of DM1-conjugated anti-TAT112 antibody, PC3.gD.NCA clone 3 cells (obtained as described in Example 14 above) were inoculated into the flank of nude mice. When the tumors

reached a mean volume of approximately 100mm3, mice were treated with DM1-conjugated anti-TAT112 antibody intravenously either once or twice per week.

The results of the above analyses demonstrated that both the unconjugated anti-TAT112 as well as the DM1-conjugated anti-TAT112 antibody were highly efficacious in reducing tumor volume in this *in vivo* model. These analyses demonstrate that anti-TAT polypeptide monoclonal antibodies are efficacious for killing tumor cells that express a TAT polypeptide of interest.

Anti-TAT118 (E16) activity:

To test the efficacy of unconjugated anti-TAT188 monoclonal antibodies, anti-TAT188 antibodies 3B5, 12B9, and 12G12 for cell killing in vitro, the following procedure was performed. PC3 cells, which express E16 endogenously, were injected into athymic nude mice (nu/nu) at the dorsal flank. On the same day as cell inoculation, antibodies were injected at 2 mg/kg intraperitoneously twice per week. The control mice were injected with PBS without antibody. Ten mice in each group were monitored for 4 weeks and tumor volume was measured twice per week. By injecting tumor cells and anti-E16 antibody simultaneously, this procedure tested the ability of the administered antibodies to inhibit tumor establishment in the xenograft mice. The results in Fig. 162 show that in this experiment, mean tumor volume was not significantly lower than for control antibody.

To test the efficacy of MC-vc-PAB-MMAE and MC-vc-PAB-MMAF conjugated anti-TAT188(E16) antibody, Colo205 cells were inoculated into the flank of athymic nude mice (nu/nu), and after approximately 1 week, when the tumor had reached a mean volume of approximately 100-200mm^3, mice were injected with conjugated antibody at 3 mg/kg intravenously once per week. The following antibody treatments were performed on the day tumor-containing mice were divided into groups: (1) Mock Control - 0.2ml or less PBS, intravenous injection once per week for 4 weeks; (2) Antibody control - anti-IL8-MC-vc-PAB-MMAE in 0.2ml or less PBS, intravenous injection once per week for 4 weeks; (3) Antibody - 3B5-MC-vc-PAB-MMAE in 0.2ml or less PBS, intravenous injection once per week for 4 weeks; (4) Antibody - 3B5-MC-vc-PAB-MMAF in 0.2ml or less PBS, intravenous injection once per week for 4 weeks. Mean tumor volume was measured and plotted versus time. The results in Fig. 163 show that MMAE- and MMAF-conjugated anti-E16 antibody significantly reduced tumor volume in this *in vivo* model. These analyses demonstrate that anti-TAT polypeptide monoclonal antibodies, such as anti-188(E16) antibodies are efficacious for killing tumor cells that express a TAT polypeptide of interest, such as the TAT188(E16) polypeptide.

EXAMPLE 16: Northern Blot Analysis

Northern blot analysis was performed essentially as described by Sambrook et al., supra. Northern blot analysis using probes derived from DNA231542, DNA231542-1, DNA231542-2 and DNA297393 evidences significant upregulation of expression in human glioma tissue as compared to normal human brain tissue.

Northern blot analysis using probes derived from DNA237637 (TAT188, E16) were also

performed and the results showed that expression in human breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver evidences significant upregulation of expression in human cancers as compared to normal human tissue of the same tissue type.

EXAMPLE 17: Modulating TAT expression by siRNA

siRNAs have proven useful as a tool in studies of modulating gene expression where traditional antagonists such as small molecules or antibodies have failed. (Shi Y., *Trends in Genetics* 19(1):9-12 (2003)). In vitro synthesized, double stranded RNAs that are 21 to 23 nucleotides in length can act as interfering RNAs (iRNAs) and can specifically inhibit gene expression (Fire A., *Trends in Genetics* 391; 806-810 (1999)). These iRNAs act by mediating degradation of their target RNAs. Since they are under 30 nuclotides in length, however they do not trigger a cell antiviral defense mechanism. Such mechanisms include interferon production, and a general shutdown of host cell protein synthesis. Practically, siRNAs can by synthesized and then cloned into DNA vectors. Such vectors can be transfected and made to express the siRNA at high levels. The high level of siRNA expression is used to "knockdown" or significantly reduce the amount of protein produced in a cell, and thus it is useful in experiments where overexpression of a protein is believed to be linked to a disorder such as cancer. The TATs are cell surface expressed tumor antigens and TAT overexpression was shown by microarray, Taqman™ and in situ hybridization analyses. While antibody binding to TATs such as TAT188 are shown herein to inhibit function of the tumor antigen and, with respect to toxin-conjugated antibodies, cause cell killing), siRNAs are useful antagonists to TAT proteins by limiting cellular production of the antigen.

Experimental Methods and Results: *TAT188 (E16)*

The following experiments demonstrate that siRNA specific for TAT188 (also referred to herein as E16) reduces RNA expression, protein production, cell surface expression, and protein function. The same or similar procedures are useful demonstrating siRNA downregulation of other TAT RNAs of the invention. Selection of a target sequence may be performed by any suitable procedure (see, for example, Amarzguioui, M. And Prydz, H., Biochem Biophys Res Comm 316:1050-1058 (2004).

A. Transient transfection of E16 siRNA reduces the expression of E16-GFP in PC3-E16-GFP cells.

The prostate cancer cell line, PC3, was used in this set of experiments as it overexpresses TAT188 endogenously. PC3 cells were transfected using double-stranded 21mer RNA oligos directed against TAT188 (siTAT188), lamin or mock transfected. These oligos are listed below.

siRNA oligos:

TAT188(E16)

The DNA sequence of the target region within the coding sequence of TAT188(E16) is shown below followed by the sequences of the sense and antisense strands of the double stranded siRNA.

AAGGAAGAGGCGCGGGAGAAG (SEQ ID NO:155), is the nucleic acid sequence within the TAT188 coding region targeted by the siRNA oligos used in the TAT188 siRNA examples disclosed herein. Other sequences may be chosen as targets according to standard practice with the field of siRNA gene silencing.

TAT188(E16) siRNA oligos:

Sense: r(GGAAGAGGCGCGGGAGAAG)TT (SEQ ID NO:156)

Antisense: r(CUUCUCCCGCGCCUCUUCC)TT (SEQ ID NO:157)

silamin siRNA oligos:

Sense: r(CUGGACUUCCAGAAGAACA)TT (SEQ ID NO:158)

Antisense: r(UGUUCUGGAAGUCCAG)TT (SEQ ID NO:159)

(see Elbashir et al., Nature 411:494-498 (2001)

Two days post-transfection and the relative level of TAT188(E16) polypeptide versus the knock-down of E16-EGFP expression was analyzed by fluorescent microscope monitoring of EGFP fluorescence. A decrease in E16 expression parallels a reduction in EGFP production and bioluminescence detected. The results in Fig. 164 demonstrate that transient transfection with siTAT188(E16) siRNA reduced E16-EGFP fusion expression. The fusion protein refered to as "GFP" in Figs. 164 and 164 refer to an E16-EGFP fusion produced by cloning the E16 gene into pEGFP-C1 vector from BDClontech).

B. Transient transfection of E16 siRNA reduces TAT188(E16) polypeptide prodcution - Reduction of E16 expression was demonstrated by reduction in E16 protein using Western blot analysis. siTAT188(siE16) RNA oligos SEQ ID NO:156 and SEQ ID NO:157 and the lamin siRNA oligos SEQ ID NO:158 and SEQ ID NO:159 used in this example were the same as shown above in part A of this example. PC3 cells transiently transfected with E16-EGFP were harvested 3 days post-transfection and a Western blot was performed on beta-tubulin band to indicate protein loading. Western blot of TAT188(E16)-GFP fusions were visualized using detectably labelled anti-GFP antibody (see Fig. 165, left lane - mock transfection; right lane, E16 siRNA transfection). The results demonstrate that the presence TAT188(E16) siRNA in a cell expressing E16-GFP fusion protein reduced the expression of E16-GFP significantly.

C. Transient transfection of E16 siRNA inhibits endogenous E16 amino acid transporter activity in PC3 cells.

The same procedure for transient transfection of TAT188(E16) siRNA and lamin siRNA (control) was performed. The test and control cells were tested for E16 activity as an amino acid transport

protein as follows using the leucine transport assay described herein above. The results in Fig. 166 show that E16 siRNA reduction in E16 RNA and protein production is consistent with reduction in the amount of E16 available to function in amino acid transport in the cell.

D. Transient transfection of E16 siRNA reduces the surface expression of E16

This example was performed as follows. Forty eight hours after transfection of siRNA, cells were analysed by FACS. Briefly, cells were incubated with primary antibodies for one hour on ice. The cells were then labeled with a secondary antibody conjugated with a fluorescent dye (such as nonlimiting examples Cy-3 or PE) and collecting data with FACScaliber™ or FACScan™. The collected data was analyzed using, as non-limiting examples, Cellquest™ (BD Clontech) or FlowJo ™ (Tree Star, Inc.). As disclosed in Example 12, herein, the three anti-E16 monoclonal antibodies, 3b5, 12B9, and 12G12, were shown to bind less on PC3 cells after transfection of TAT188(E16) siRNA consistent with reduced expression of the E16 protein in the presence of targeted siRNA (see Fig. 156).

E. Transient transfection of TAT188(E16) siRNA inhibits PC3 cell proliferation.

Proliferation of PC3 cells transfected with siRNA targeting TAT188 can be assayed according to the following commercial protocol (Promega Corp. Technical Bulletin TB288; Mendoza et al. (2202) Cancer Res. 62:5485-5488). For example, PC3 cells were plated at $3.5x10^5$ cells/well in 6 well plates. The next day, double stranded RNA (dsRNA, SEQ ID NO:155 and SEQ ID NO:156) or lamin control siRNA (using siRNA oligos SEQ ID NO:157 and SEQ ID NO:158) at a final concentration of 100 nM was used to transfect the cells using Lipofectamin2000 (Invitrogen). After 6 hrs of transfection, cells were reseeded into 96 well plates, and cell viability was measured using Cell Counting Kit-8™ (Dojindo) at the time points indicated in Fig. 167, where RLU = relative luminescence units. Other cells that may be assayed by this method if they express TAT188(either endogenousl y or after transfection with nucleic acid encoding TAT188) include without limitation SKBR-3, BT474, MCF7 or MDA-MB-468 demonstrating that reduction in E16 expression in breast cancer cells results in cell killing and that breast cancer, as well as prostate, colorectal and other types of cancer, are desirable targets for treatment using E16 expression knockdown.

These data shows that siRNAs can "knockdown" or reduce the expression of TAT proteins such as TAT188 (E16) protein. The reduction is shown by reduction of RNA expression (in *in situ* hybridization assays), reduced protein production (as shown in Western blot analyses), reduced cell surface expression of TAT188 (E16) (as shown by FACS analyses and internalization assays), and reduced function (as shown by reduction in amino acid transport). siRNA reduction of TAT proteins, such as TAT188 (E16) over a period of time results in reduction of cell proliferation in cancer cell lines. Therefore, siRNA is useful in reducing TAT proteins such as TAT188 (E16). Reduction in the level of TAT proteins, such as TAT188 (E16), causes a reduction in the proliferation of cancer cell lines. Since cancer is a cell proliferative disorder, these results show that antagonists of TAT polypeptides, such as TAT188 (E16) would be useful in the reduction of cancer cell growth. Reduction of cancer cell growth would be useful in the alleviation of cancer in mammals.

Example 18: Deposit of Materials

The following hybridoma cell line has been deposited with the American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209 USA (ATCC):

Hybridoma/Antibody Designation	ATCC No.	Deposit Date
3B5.1	PTA-6193	September 8, 2004
12B9.1	PTA-6194	September 8, 2004
12G12.1	PTA-6195	September 8, 2004

This deposit was made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture for 30 years from the date of deposit. The cell line will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures (a) that access to the culture will be available during pendency of the patent application to one determined by the Commissioner to be entitled thereto under 37 CFR §1.14 and 35 USC §122, and (b) that all restrictions on the availability to the public of the culture so deposited will be irrevocably removed upon the granting of the patent.

The assignee of the present application has agreed that if the culture on deposit should die or be lost or destroyed when cultivated under suitable conditions, it will be promptly replaced on notification with a viable specimen of the same culture. Availability of the deposited cell line is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. An isolated antibody that competes with binding to the epitopes of a polypeptide having amino acid sequence of SEQ ID NO:115 bound by an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195).

2. The antibody of Claim 1 which is conjugated to a growth inhibitory agent.

3. The antibody of Claim 1 which is conjugated to a cytotoxic agent.

4. The antibody of Claim 3, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

5. The antibody of Claim 3, wherein the cytotoxic agent is a toxin.

6. The antibody of Claim 5, wherein the toxin is selected from the group consisting of auristatin, maytansinoid and calicheamicin.

7. The antibody of Claim 5, wherein the toxin is a maytansinoid.

8. The antibody of Claim 1 which is detectably labeled.

9. The antibody of Claim 5, wherein the toxin is selected from the group consisting of MMAE (mono-methyl auristatin E), MMAF, AEVB (auristatin E valeryl benzylhydrazone), and AFP (Auristatin F phenylene diamine).

10. The antibody of claim 5, wherein the toxin is covalently attached to the antibody by a linker.

11. The antibody of claim 10, wherein the linker is selected from the group consisting of maleimidocaproyl (MC), valine-citrulline (val-cit, vc), citrulline (2-amino-5-ureido pentanoic acid), PAB (p-aminobenzylcarbamoyl), Me (N-methyl-valine citrulline), MC(PEG)6-OH (maleimidocaproyl-polyethylene glycol), SPP (N-Succinimidyl 4-(2-pyridylthio) pentanoate), SMCC (N-Succinimidyl 4-(N-maleimidomethyl) cyclohexane-1 carboxylate), and MC-vc-PAB.

12. The antibody of claim 9, wherein the toxin is MMAE.

13. The antibody of claim 9, wherein the toxin is MMAF.

14. The antibody of claim 12 or 13, wherein the linker is MC-vc-PAB.

15. An isolated antibody having the biological activity of an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195), wherein the biological activity is inhibition of cell proliferation or promotion of cell death in a cell expressing a polypeptide having amino acid sequence of SEQ ID NO:115. .

16. An isolated antibody of claim 15 comprising in a corresponding complementary determining region (CDR) an amino acid sequence having at least 95%, 96%, 97%, 98%, 99% or 100% of the amino acid sequence of –6 of the CDR(s) of the antibody produced by a hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1

(ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195)

17. The antibody of claim 15, wherein the cell is a cancer cell.

18. The antibody of claim 17, wherein the cancer cell is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

19. A composition of matter comprising: (a) the antibody of Claim 1 in combination with a carrier.

20. The composition of matter of Claim 19, wherein said carrier is a pharmaceutically acceptable carrier.

21. An article of manufacture:

(a) a container; and

(b) the composition of matter of Claim 19 contained within said container.

22. The article of manufacture of Claim 21, further comprising a label affixed to said container, or a package insert included with said container, referring to the use of said composition of matter for the therapeutic treatment of or the diagnostic detection of a cancer.

Sequence Listing

<110> GENENTECH, INC.

<120> Compositions and Methods for the Diagnosis and Treatment of Tumor

<130> P5001R1P1 PCT

<140> PCT/US2005/031798
<141> 2005-07-09

<150> US 10/936,626
<151> 2004-09-08

<150> US 10/177,488
<151> 2002-06-19

<150> US 60/299,500
<151> 2001-06-20

<150> US 60/301,880
<151> 2001-06-29

<150> US 60/323,268
<151> 2001-09-18

<160> 159

<210> 1
<211> 1039
<212> DNA
<213> Homo sapiens

<400> 1
ccaagaattc ggcacgaggg tgacccggac agctgtcctc tctgacacca 50

ccccggcctg cctctttgtt gccatgagag ctgcctacct cttcctgcta 100

ttcctgcctg caggcttgct ggctcagggc cagtatgacc tggacccgct 150

gccgccgttc cctgaccacg tccagtacac ccactatagc gaccagatcg 200

acaacccaga ctactatgat tatcaagagg tgactcctcg gccctccgag 250

gaacagttcc agttccagtc ccagcagcaa gtccaacagg aagtcatccc 300

agccccaacc ccagaaccag gaaatgcaga gctggagccc acagagcctg 350

ggcctcttga ctgccgtgag gaacagtacc cgtgcacccg cctctactcc 400

atacacaggc cttgcaaaca gtgtctcaac gaggtctgct tctacagcct 450

ccgccgtgtg tacgtcatta acaaggagat ctgtgttcgt acagtgtgtg 500

cccacgagga gctcctccga gctgacctct gtcgggacaa gttctccaaa 550

tgtggcgtga tggccagcag cggcctgtgc caatccgtgg cggcctcctg 600

tgccaggagc tgtgggagct gctagggtgg tgctggcatc ctgagtcctg 650

gccctcctgg gatctggggc cctcgggccc tgcctgacct ggtgcttttt 700

tccccatccc catgttcctt ttattctgta aaaagttagt ggactgcagc 750

cctgggggtt gcaggctgcg gtgcctcagg cccctccttc agcctgtggc 800

cacctctggg gcacgatggg ggctccccac tgcccagtct gcccctcggg 850

ttgggggagt atcccaggcc tctctgtggg acctgggccc ctgacgggcc 900

ttctcagccc gttttgagga cagacagtcc cccgaggtag gctacatccc 950

cccaccccag ctggtctgct tggatttcct acagcccccg tgggcatgga 1000

ccacctttat tttatacaaa attaaaaaca agtttttac 1039

<210> 2
<211> 5062
<212> DNA
<213> Homo sapiens

<400> 2
cactaacgct cttcctagtc cccgggccaa ctcggacagt ttgctcattt 50

attgcaacgg tcaaggctgg cttgtgccag aacggcgcgc gcgcgacgca 100

cgcacacaca cgggggggaaa ctttttttaaa aatgaaaggc tagaagagct 150

cagcggcggc gcgggccgtg cgcgagggct ccggagctga ctcgccgagg 200

caggaaatcc ctccggtcgc gacgcccggc cccgctcggc gcccgcgtgg 250

gatggtgcag cgctcgccgc cgggcccgag agctgctgca ctgaaggccg 300

gcgacgatgg cagcgcgccc gctgcccgtg tcccccgccc gcgccctcct 350

gctcgccctg gccggtgctc tgctcgcgcc ctgcgaggcc cgaggggtga 400

gcttatggaa cgaaggaaga gctgatgaag ttgtcagtgc ctctgttcgg 450

agtggggacc tctggatccc agtgaagagc ttcgactcca agaatcatcc 500

agaagtgctg aatattcgac tacaacggga aagcaaagaa ctgatcataa 550

atctggaaag aaatgaaggt ctcattgcca gcagtttcac ggaaacccac 600

tatctgcaag acggtactga tgtctccctc gctcgaaatt acacggtaat 650

tctgggtcac tgttactacc atggacatgt acggggatat tctgattcag 700

cagtcagtct cagcacgtgt tctggtctca ggggacttat tgtgtttgaa 750

aatgaaagct atgtcttaga accaatgaaa agtgcaacca acagatacaa 800

actcttccca gcgaagaagc tgaaaagcgt ccggggatca tgtggatcac 850

atcacaacac accaaacctc gctgcaaaga atgtgtttcc accaccctct 900

cagacatggg caagaaggca taaaagagag accctcaagg caactaagta 950

tgtggagctg gtgatcgtgg cagacaaccg agagtttcag aggcaaggaa 1000

aagatctgga aaaagttaag cagcgattaa tagagattgc taatcacgtt 1050

gacaagtttt acagaccact gaacattcgg atcgtgttgg taggcgtgga 1100

agtgtggaat gacatggaca aatgctctgt aagtcaggac ccattcacca 1150

gcctccatga atttctggac tggaggaaga tgaagcttct acctcgcaaa 1200

tcccatgaca atgcgcagct tgtcagtggg gtttatttcc aagggaccac 1250

catcggcatg gccccaatca tgagcatgtg cacggcagac cagtctgggg 1300

gaattgtcat ggaccattca gacaatcccc ttggtgcagc cgtgaccctg 1350

gcacatgagc tgggccacaa tttcgggatg aatcatgaca cactggacag 1400

gggctgtagc tgtcaaatgg cggttgagaa aggaggctgc atcatgaacg 1450

cttccaccgg gtacccattt cccatggtgt tcagcagttg cagcaggaag 1500

gacttggaga ccagcctgga gaaaggaatg ggggtgtgcc tgtttaacct 1550

gccggaagtc agggagtctt tcgggggcca gaagtgtggg aacagatttg 1600

tggaagaagg agaggagtgt gactgtgggg agccagagga atgtatgaat 1650

cgctgctgca atgccaccac ctgtaccctg aagccggacg ctgtgtgcgc 1700

acatgggctg tgctgtgaag actgccagct gaagcctgca ggaacagcgt 1750

gcagggactc cagcaactcc tgtgacctcc cagagttctg cacaggggcc 1800

agccctcact gcccagccaa cgtgtacctg cacgatgggc actcatgtca 1850

ggatgtggac ggctactgct acaatggcat ctgccagact cacgagcagc 1900

agtgtgtcac actctgggga ccaggtgcta aacctgcccc tgggatctgc 1950

tttgagagag tcaattctgc aggtgatcct tatggcaact gtggcaaagt 2000

ctcgaagagt tcctttgcca aatgcgagat gagagatgct aaatgtggaa 2050

aaatccagtg tcaaggaggt gccagccggc cagtcattgg taccaatgcc 2100

gtttccatag aaacaaacat ccccctgcag caaggaggcc ggattctgtg 2150

ccgggggacc cacgtgtact tgggcgatga catgccggac ccagggcttg 2200

tgcttgcagg cacaaagtgt gcagatggaa aaatctgcct gaatcgtcaa 2250

tgtcaaaata ttagtgtctt tggggttcac gagtgtgcaa tgcagtgcca 2300

cggcagaggg gtgtgcaaca acaggaagaa ctgccactgc gaggcccact 2350

gggcacctcc cttctgtgac aagtttggct ttggaggaag cacagacagc 2400

ggccccatcc ggcaagcaga taaccaaggt ttaaccatag gaattctggt 2450

gaccatcctg tgtcttcttg ctgccggatt tgtggtttat ctcaaaagga 2500

agaccttgat acgactgctg tttacaaata agaagaccac cattgaaaaa 2550

ctaaggtgtg tgcgcccttc ccggccaccc cgtggcttcc aaccctgtca 2600

ggctcacctc ggccaccttg gaaaaggcct gatgaggaag ccgccagatt 2650

cctacccacc gaaggacaat cccaggagat tgctgcagtg tcagaatgtt 2700

gacatcagca gacccctcaa cggcctgaat gtccctcagc cccagtcaac 2750

tcagcgagtg cttcctcccc tccaccgggc cccacgtgca cctagcgtcc 2800

ctgccagacc cctgccagcc aagcctgcac ttaggcaggc ccaggggacc 2850

tgtaagccaa acccccctca gaagcctctg cctgcagatc ctctggccag 2900

aacaactcgg ctcactcatg ccttggccag gaccccagga caatgggaga 2950

ctgggctccg cctggcaccc ctcagacctg ctccacaata tccacaccaa 3000

gtgcccagat ccacccacac cgcctatatt aagtgagaag ccgacacctt 3050

ttttcaacag tgaagacaga agtttgcact atctttcagc tccagttgga 3100

gttttttgta ccaactttta ggatttttt taatgtttaa aacatcatta 3150

ctataagaac tttgagctac tgccgtcagt gctgtgctgt gctatggtgc 3200

tctgtctact tgcacaggta cttgtaaatt attaatttat gcagaatgtt 3250

gattacagtg cagtgcgctg tagtaggcat ttttaccatc actgagtttt 3300

ccatggcagg aaggcttgtt gtgcttttag tattttagtg aacttgaaat 3350

atcctgcttg atgggattct ggacaggatg tgtttgcttt ctgatcaagg 3400

ccttattgga aagcagtccc ccaactaccc ccagctgtgc ttatggtacc 3450

agatgcagct caagagatcc caagtagaat ctcagttgat tttctggatt 3500

ccccatctca ggccagagcc aaggggcttc aggtccaggc tgtgtttggc 3550

tttcagggag gccctgtgcc ccttgacaac tggcaggcag gctcccaggg 3600

acacctggga gaaatctggc ttctggccag gaagctttgg tgagaacctg 3650

ggttgcagac aggaatctta aggtgtagcc acaccaggat agagactgga 3700

acactagaca agccagaact tgaccctgag ctgaccagcc gtgagcatgt 3750

ttggaagggg tctgtagtgt cactcaaggc ggtgcttgat agaaatgcca 3800

agcacttctt tttctcgctg tcctttctag agcactgcca ccagtaggtt 3850

atttagcttg ggaaaggtgg tgtttctgta agaaacctac tgcccaggca 3900

ctgcaaaccg ccacctccct atactgcttg gagctgagca aatcaccaca 3950

aactgtaata caatgatcct gtattcagac agatgaggac tttccatggg 4000

accacaacta ttttcagatg tgaaccatta accagatcta gtcaatcaag 4050

tctgtttact gcaaggttca acttattaac aattaggcag actctttatg 4100

cttgcaaaaa ctacaaccaa tggaatgtga tgttcatggg tatagttcat 4150

gtctgctatc attattcgta gatattggac aaagaacctt ctctatgggg 4200

catcctcttt ttccaacttg gctgcaggaa tctttaaaag atgcttttaa 4250

cagagtctga acctatttct taaacacttg caacctacct gttgagcatc 4300

acagaatgtg ataaggaaat caacttgctt atcaacttcc taaatattat 4350

gagatgtggc ttgggcagca tccccttgaa ctcttcactc ttcaaatgcc 4400

tgactaggga gccatgtttc acaaggtctt taaagtgact aatggcatga 4450

gaaatacaaa aatactcaga taaggtaaaa tgccatgatg cctctgtctt 4500

ctggactggt tttcacatta gaagacaatt gacaacagtt acataattca 4550

ctctgagtgt tttatgagaa agccttcttt tggggtcaac agttttccta 4600

tgctttgaaa cagaaaaata tgtaccaaga atcttggttt gccttccaga 4650

aaacaaaact gcatttcact ttcccggtgt tccccactgt atctaggcaa 4700

catagtattc atgactatgg ataaactaaa cacgtgacac aaacacacac 4750

aaaagggaac ccagctctaa tacattccaa ctcgtatagc atgcatctgt 4800

ttattctata gttattaagt tctttaaaat gtaaagccat gctggaaaat 4850

aatactgctg agatacatac agaattactg taactgatta cacttggtaa 4900

ttgtactaaa gccaaacata tatatactat taaaaaggtt tacagaattt 4950

tatggtgcat tacgtgggca ttgtcttttt agatgcccaa atccttagat 5000

ctggcatgtt agcccttcct ccaattataa gaggatatga accaaaaaaa 5050

aaaaaaaaaa aa 5062

<210> 3
<211> 1150
<212> DNA
<213> Homo sapiens

<400> 3
ggagtttcaa agaagcgcta gtaaggtctc tgagatcctt gcactagcta 50

catcctcagg gtaggaggaa gatggcttcc agaagcatgc ggctgctcct 100

attgctgagc tgcctggcca aaacaggagt cctgggtgat atcatcatga 150

gacccagctg tgctcctgga tggttttacc acaagtccaa ttgctatggt 200

tacttcagga agctgaggaa ctggtctgat gccgagctcg agtgtcagtc 250

ttacggaaac ggagcccacc tggcatctat cctgagttta aaggaagcca 300

gcaccatagc agagtacata agtggctatc agagaagcca gccgatatgg 350

attggcctgc acgacccaca gaagaggcag cagtggcagt ggattgatgg 400

ggccatgtat ctgtacagat cctggtctgg caagtccatg ggtgggaaca 450

agcactgtgc tgagattagc tccaataaca acttttaac ttggagcagc 500

aacgaatgca acaagcgcca acacttcctg tgcaagtacc gaccatagag 550

caagaatcaa gattctgcta actcctgcac agccccgtcc tcttcctttc 600

tgctagcctg gctaaatctg ctcattattt cagaggggaa acctagcaaa 650

ctaagagtga taagggccct actacactgg ctttttagg cttagagaca 700

gaaactttag cattggccca gtagtggctt ctagctctaa atgtttgccc 750

cgccatccct ttccacagta tccttcttcc ctcctcccct gtctctggct 800

gtctcgagca gtctagaaga gtgcatctcc agcctatgaa acagctgggt 850

ctttggccat aagaagtaaa gatttgaaga cagaaggaag aaactcagga 900

gtaagcttct agacccccttc agcttctaca cccttctgcc ctctctccat 950

tgcctgcacc ccaccccagc cactcaactc ctgcttgttt ttcctttggc 1000

cataggaagg tttaccagta gaatccttgc taggttgatg tgggccatac 1050

attcctttaa taaaccattg tgtacataaa aaaaaaaaaa aaaaaaaaaa 1100

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1150

<210> 4
<211> 474
<212> DNA
<213> Homo sapiens

<400> 4
ggggagcaga gaggaggcaa tggccaccat ggagaacaag gtgatctgcg 50

ccctggtcct ggtgtccatg ctggccctcg gcaccctggc cgaggcccag 100

acagagacgt gtacagtggc cccccgtgaa agacagaatt gtggttttcc 150

tggtgtcacg ccctcccagt gtgcaaataa gggctgctgt ttcgacgaca 200

ccgttcgtgg ggtcccctgg tgcttctatc ctaataccat cgacgtccct 250

ccagaagagg agtgtgaatt ttagacactt ctgcagggat ctgcctgcat 300

cctgacgcgg tgccatcccc agcacggtga ttagtcccag agctcggctg 350

ccacctccac cggacacctc agacacgctt ctgcagctgt gcctcggctc 400

acaacacaga ttgactgctc tgactttgac tactcaaaat tggcctaaaa 450

attaaaagag ctcgatatta aaaa 474

<210> 5

<211> 3226
<212> DNA
<213> Homo sapiens

<400> 5
aatccatctg agaatatgct gccacaaata ccctttttgc tgctagtatc 50

cttgaacttg gttcatggag tgttttacgc tgaacgatac caaacgccca 100

caggcataaa aggcccacta cccaacacca agacacagtt cttcattccc 150

tacaccataa agagtaaagg tatagcagta agaggagagc aaggtactcc 200

tggtccacca ggccctgctg gacctcgagg gcacccaggt ccttctggac 250

caccaggaaa accaggctac ggaagtcctg gactccaagg agagccaggg 300

ttgccaggac caccgggacc atcagctgta gggaaaccag gtgtgccagg 350

actcccagga aaaccaggag agagaggacc atatggacca aaaggagatg 400

ttggaccagc tggcctacca ggaccccggg gcccaccagg accacctgga 450

atccctggac cggctggaat ttctgtgcca ggaaaacctg gacaacaggg 500

acccacagga gccccaggac ccaggggctt tcctggagaa aagggtgcac 550

caggagtccc tggtatgaat ggacagaaag gggaaatggg atatggtgct 600

cctggtcgtc caggtgagag gggtcttcca ggccctcagg gtcccacagg 650

accatctggc cctcctggag tgggaaaaag aggtgaaaat ggggttccag 700

gacagccagg catcaaaggt gatagaggtt ttccgggaga aatgggacca 750

attggcccac caggtcccca aggccctcct ggggaacgag ggccagaagg 800

cattggaaag ccaggagctg ctggagcccc aggccagcca gggattccag 850

gaacaaaagg tctccctggg gctccaggaa tagctgggcc cccagggcct 900

cctggctttg ggaaaccagg cttgccaggc ctgaagggag aaagaggacc 950

tgctggcctt cctgggggtc caggtgccaa aggggaacaa gggccagcag 1000

gtcttcctgg gaagccaggt ctgactggac cccctgggaa tatgggaccc 1050

caaggaccaa aaggcatccc gggtagccat ggtctcccag gccctaaagg 1100

tgagacaggg ccagctgggc ctgcaggata ccctggggct aagggtgaaa 1150

ggggttcccc tgggtcagat ggaaaaccag ggtacccagg aaaaccaggt 1200

ctcgatggtc ctaagggtaa cccagggtta ccaggtccaa aaggtgatcc 1250

tggagttgga ggacctcctg gtctcccagg ccctgtgggc ccagcaggag 1300

caaagggaat gcccggacac aatggagagg ctggcccaag aggtgcccct 1350

ggaataccag gtactagagg ccctattggg ccaccaggca ttccaggatt 1400

ccctgggtct aaaggggatc caggaagtcc cggtcctcct ggcccagctg 1450

gcatagcaac taagggcctc aatggaccca ccgggccacc agggcctcca 1500

ggtccaagag gcccctctgg agagcctggt cttccagggc cccctgggcc 1550

tccaggccca ccaggtcaag cagtcatgcc tgagggtttt ataaaggcag 1600

gccaaaggcc cagtctttct gggacccctc ttgttagtgc caaccagggg 1650

gtaacaggaa tgcctgtgtc tgcttttact gttattctct ccaaagctta 1700

cccagcaata ggaactccca taccatttga taaaattttg tataacaggc 1750

aacagcatta tgacccaagg actggaatct ttacttgtca gataccagga 1800

atatactatt tttcatacca cgtgcatgtg aaagggactc atgtttgggt 1850

aggcctgtat aagaatggca cccctgtaat gtacacctat gatgaataca 1900

ccaaaggcta cctggatcag gcttcaggga gtgccatcat cgatctcaca 1950

gaaaatgacc aggtgtggct ccagcttccc aatgccgagt caaatggcct 2000

atactcctct gagtatgtcc actcctcttt ctcaggattc ctagtggctc 2050

caatgtgagt acaccccaca gagctaatct aaatcttgtg ctagaaaaag 2100

cattctctaa ctctacccca ccctacaaaa tgcatatgga ggtaggctga 2150

aaagaatgta atttttattt tctgaaatac agatttgagc tatcagacca 2200

acaaaccttc cccctgaaaa gtgagcagca acgtaaaaac gtatgtgaag 2250

cctctcttga atttctagtt agcaatctta aggctcttta aggttttctc 2300

caatattaaa aaatatcacc aaagaagtcc tgctatgtta aaaacaaaca 2350

acaaaaaaca aagcaacaaa aaaaaaaatt aaaaaaaaaa acagaaatag 2400

agctctaagt tatgtgaaat ttgatttgag aaactcggca tttccttttt 2450

aaaaaagcct gtttctaact atgaatatga gaacttctag gaaacatcca 2500

ggaggtatca tataactttg tagaacttaa atacttgaat attcaaattt 2550

aaaagacact gtatccccta aaatatttct gatggtgcac tactctgagg 2600

cctgtatggc ccctttcatc aatatctatt caaatataca ggtgcatata 2650

tacttgttaa agctcttata taaaaaagcc ccaaaatatt gaagttcatc 2700

tgaaatgcaa ggtgctttca tcaatgaacc ttttcaaaac ttttctatga 2750

ttgcagagaa gctttttata tacccagcat aacttggaaa caggtatctg 2800

acctattctt atttagttaa cacaagtgtg attaatttga tttctttaat 2850

tccttattga atcttatgtg atatgatttt ctggatttac agaacattag 2900

cacatgtacc ttgtgcctcc cattcaagtg aagttataat ttacactgag 2950

ggtttcaaaa ttcgactaga agtggagata tattatttat ttatgcactg 3000

tactgtattt ttatattgct gtttaaaact tttaagctgt gcctcactta 3050

ttaaagcaca aaatgtttta cctactcctt atttacgaca caataaaata 3100

acatcaatag atttttaggc tgaattaatt tgaaagcagc aatttgctgt 3150

tctcaaccat tctttcaagg cttttcattc gacacaataa aataacatca 3200

atagattttt agggatgggt ggcttt 3226

<210> 6
<211> 1750
<212> DNA
<213> Homo sapiens

<400> 6
catgccgctg ccgccgctgc tgctgttgct cctggcggcg ccttggggac 50

gggcagttcc ctgtgtctct ggtggtttgc ctaaacctgc aaacatcacc 100

ttcttatcca tcaacatgaa gaatgtccta caatggactc caccagaggg 150

tcttcaagga gttaaagtta cttacactgt gcagtatttc atatatgggc 200

aaaagaaatg gctgaataaa tcagaatgca gaaatatcaa tagaacctac 250

tgtgatcttt ctgctgaaac ttctgactac gaacaccagt attatgccaa 300

agttaaggcc atttggggaa caaagtgttc caaatgggct gaaagtggac 350

ggttctatcc ttttttagaa acacaaattg gcccaccaga ggtggcactg 400

actacagatg agaagtccat ttctgttgtc ctgacagctc cagagaagtg 450

gaagagaaat ccagaagacc ttcctgtttc catgcaacaa atatactcca 500

atctgaagta taacgtgtct gtgttgaata ctaaatcaaa cagaacgtgg 550

tcccagtgtg tgaccaacca cacgctggtg ctcacctggc tggagccgaa 600

cactctttac tgcgtacacg tggagtcctt cgtcccaggg ccccctcgcc 650

gtgctcagcc ttctgagaag cagtgtgcca ggactttgaa agatcaatca 700

tcagagttca aggctaaaat catcttctgg tatgttttgc ccatatctat 750

taccgtgttt ctttttttctg tgatgggcta ttccatctac cgatatatcc 800

acgttggcaa agagaaacac ccagcaaatt tgattttgat ttatggaaat 850

gaatttgaca aaagattctt tgtgcctgct gaaaaaatcg tgattaactt 900

tatcaccctc aatatctcgg atgattctaa aatttctcat caggatatga 950

gtttactggg aaaaagcagt gatgtatcca gccttaatga tcctcagccc 1000

agcgggaacc tgaggccccc tcaggaggaa gaggaggtga aacatttagg 1050

gtatgcttcg catttgatgg aaatttttg tgactctgaa gaaaacacgg 1100

aaggtacttc tctcacccag caagagtccc tcagcagaac aataccccg 1150

gataaaacag tcattgaata tgaatatgat gtcagaacca ctgacatttg 1200

tgcggggcct gaagagcagg agctcagttt gcaggaggag gtgtccacac 1250

aaggaacatt attggagtcg caggcagcgt tggcagtctt gggcccgcaa 1300

acgttacagt actcatacac ccctcagctc caagacttag acccccctggc 1350

gcaggagcac acagactcgg aggaggggcc ggaggaagag ccatcgacga 1400

ccctggtcga ctgggatccc caaactggca ggctgtgtat tccttcgctg 1450

tccagcttcg accaggattc agagggctgc gagccttctg agggggatgg 1500

gctcggagag gagggtcttc tatctagact ctatgaggag ccggctccag 1550

acaggccacc aggagaaaat gaaacctatc tcatgcaatt catggaggaa 1600

tgggggttat atgtgcagat ggaaaactga tgccaacact tccttttgcc 1650

ttttgtttcc tgtgcaaaca agtgagtcac ccctttgatc ccagccataa 1700

agtacctggg atgaaagaag ttttttccag tttgtcagtg tctgtgagaa 1750

<210> 7
<211> 1841
<212> DNA
<213> Homo sapiens

<400> 7
agcgggcgtg ggactgagca gtctgctgcc ccccgacatg tgacccagcc 50

ccgccgccca tgcgggctcc cggccgcccg gccctgcggc cgctgccgct 100

gccgccgctg ctgctgttgc tcctggcggc gccttgggga cgggcagttc 150

cctgtgtctc tggtggtttg cctaaacctg caaacatcac cttcttatcc 200

atcaacatga agaatgtcct acaatggact ccaccagagg gtcttcaagg 250

agttaaagtt acttacactg tgcagtattt catatatggg caaaagaaat 300

ggctgaataa atcagaatgc agaaatatca atagaaccta ctgtgatctt 350

tctgctgaaa cttctgacta cgaacaccag tattatgcca aagttaaggc 400

catttgggga acaaagtgtt ccaaatgggc tgaaagtgga cggttctatc 450

ctttttaga aacacaaatt ggcccaccag aggtggcact gactacagat 500

gagaagtcca tttctgttgt cctgacagct ccagagaagt ggaagagaaa 550

tccagaagac cttcctgttt ccatgcaaca aatatactcc aatctgaagt 600

ataacgtgtc tgtgttgaat actaaatcaa acagaacgtg gtcccagtgt 650

gtgaccaacc acacgctggt gctcacctgg ctggagccga acactcttta 700

ctgcgtacac gtggagtcct tcgtcccagg gccccctcgc cgtgctcagc 750

cttctgagaa gcagtgtgcc aggactttga aagatcaatc atcagagttc 800

aaggctaaaa tcatcttctg gtatgttttg cccatatcta ttaccgtgtt 850

tctttttct gtgatgggct attccatcta ccgatatatc cacgttggca 900

aagagaaaca cccagcaaat ttgattttga tttatggaaa tgaatttgac 950

aaaagattct ttgtgcctgc tgaaaaaatc gtgattaact ttatcaccct 1000

caatatctcg gatgattcta aaatttctca tcaggatatg agtttactgg 1050

gaaaaagcag tgatgtatcc agccttaatg atcctcagcc cagcgggaac 1100

ctgaggcccc ctcaggagga agaggaggtg aaacatttag ggtatgcttc 1150

gcatttgatg gaaatttttt gtgactctga agaaaacacg gaaggtactt 1200

ctctcaccca gcaagagtcc ctcagcagaa caatacccc ggataaaaca 1250

gtcattgaat atgaatatga tgtcagaacc actgacattt gtgcggggcc 1300

tgaagagcag gagctcagtt tgcaggagga ggtgtccaca caaggaacat 1350

tattggagtc gcaggcagcg ttggcagtct tgggcccgca aacgttacag 1400

tactcataca cccctcagct ccaagactta gacccccctgg cgcaggagca 1450

cacagactcg gaggaggggc cggaggaaga gccatcgacg accctggtcg 1500

actgggatcc ccaaactggc aggctgtgta ttccttcgct gtccagcttc 1550

gaccaggatt cagagggctg cgagccttct gaggggggatg ggctcggaga 1600

ggagggtctt ctatctagac tctatgagga gccggctcca gacaggccac 1650

caggagaaaa tgaaacctat ctcatgcaat tcatggagga atgggggtta 1700

tatgtgcaga tggaaaactg atgccaacac ttccttttgc cttttgtttc 1750

ctgtgcaaac aagtgagtca ccccttttgat cccagccata aagtacctgg 1800

gatgaaagaa gttttttcca gtttgtcagt gtctgtgaga a 1841

<210> 8
<211> 536
<212> DNA
<213> Homo sapiens

<400> 8
gcttggtcta gtgcccattt actctggact ccggatggct gccgcacgcc 50

tctgcctctc cctgctgctc ctgtccacct gcgtggctct gttactacag 100

ccactgctgg gtgcccaggg agccccactg gagccagtgt acccagggga 150

caatgccaca ccagagcaga tggcccagta tgcagctgat ctccgtagat 200

acatcaacat gctgaccagg cctaggtatg ggaaaagaca caaagaggac 250

acgctggcct tctcggagtg ggggtccccg catgctgctg tccccaggga 300

gctcagcccg ctggacttat aatgccacct tctgtctcct acgactccat 350

gagcagcgcc agcccagctc tcccctctgc acccttggct ctggccaaag 400

cttgctccct gctcccacac aggctcaata aagcaagtca aagccaaaaa 450

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 500

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 536

<210> 9
<211> 2588
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 1940
<223> unknown base

<400> 9
gaaggacagc agggccaaca gtcacagcag ccctgaccag agcattcctg 50

gagctcaagc tcctctacaa agaggtggac agagaagaca gcagagacca 100

tgggaccccc ctcagcccct ccctgcagat tgcatgtccc ctggaaggag 150

gtcctgctca cagcctcact tctaaccttc tggaacccac ccaccactgc 200

caagctcact attgaatcca cgccgttcaa tgtcgcagag gggaaggagg 250

ttcttctact cgcccacaac ctgccccaga atcgtattgg ttacagctgg 300

tacaaaggcg aaagagtgga tggcaacagt ctaattgtag gatatgtaat 350

aggaactcaa caagctaccc cagggcccgc atacagtggt cgagagacaa 400

tataccccaa tgcatccctg ctgatccaga acgtcaccca gaatgacaca 450

ggattctata ccctacaagt cataaagtca gatcttgtga atgaagaagc 500

aaccggacag ttccatgtat acccggagct gcccaagccc tccatctcca 550

gcaacaactc caaccccgtg gaggacaagg atgctgtggc cttcacctgt 600

gaacctgagg ttcagaacac aacctacctg tggtgggtaa atggtcagag 650

cctcccggtc agtcccaggc tgcagctgtc caatggcaac atgaccctca 700

ctctactcag cgtcaaaagg aacgatgcag gatcctatga atgtgaaata 750

cagaacccag cgagtgccaa ccgcagtgac ccagtcaccc tgaatgtcct 800

ctatggccca gatggcccca ccatttcccc ctcaaaggcc aattaccgtc 850

caggggaaaa tctgaacctc tcctgccacg cagcctctaa cccacctgca 900

cagtactctt ggtttatcaa tgggacgttc cagcaatcca cacaagagct 950

ctttatcccc aacatcactg tgaataatag cggatcctat atgtgccaag 1000

cccataactc agccactggc ctcaatagga ccacagtcac gatgatcaca 1050

gtctctggaa gtgctcctgt cctctcagct gtggccaccg tcggcatcac 1100

gattggagtg ctggccaggg tggctctgat atagcagccc tggtgtattt 1150

tcgatatttc aggaagactg gcagattgga ccagaccctg aattcttcta 1200

gctcctccaa tcccatttta tcccatggaa ccactaaaaa caaggtctgc 1250

tctgctcctg aagccctata tgctggagat ggacaactca atgaaaattt 1300

aaagggaaaa ccctcaggcc tgaggtgtgt gccactcaga gacttcacct 1350

aactagagac aggcaaactg caaaccatgg tgagaaattg acgacttcac 1400

actatggaca gcttttccca agatgtcaaa acaagactcc tcatcatgat 1450

aaggctctta ccccctttta atttgtcctt gcttatgcct gcctctttcg 1500

cttggcagga tgatgctgtc attagtattt cacaagaagt agcttcagag 1550

ggtaacttaa cagagtatca gatctatctt gtcaatccca acgttttaca 1600

taaaataaga gatcctttag tgcacccagt gactgacatt agcagcatct 1650

ttaacacagc cgtgtgttca aatgtacagt ggtccttttc agagttggac 1700

ttctagactc acctgttctc actccctgtt ttaattcaac ccagccatgc 1750

aatgccaaat aatagaattg ctccctacca gctgaacagg gaggagtctg 1800

tgcagtttct gacacttgtt gttgaacatg gctaaataca atgggtatcg 1850

ctgagactaa gttgtagaaa ttaacaaatg tgctgcttgg ttaaaatggc 1900

tacactcatc tgactcattc tttattctat tttagttggn ttgtatcttg 1950

cctaaggtgc gtagtccaac tcttggtatt accctcctaa tagtcatact 2000

agtagtcata ctccctggtg tagtgtattc tctaaaagct ttaaatgtct 2050

gcatgcagcc agccatcaaa tagtgaatgg tctctctttg gctggaatta 2100

caaaactcag agaaatgtgt catcaggaga acatcataac ccatgaagga 2150

taaaagcccc aaatggtggt aactgataat agcactaatg ctttaagatt 2200

tggtcacact ctcacctagg tgagcgcatt gagccagtgg tgcataaatg 2250

ctacatactc caactgaaat gttaaggaag aagatagatc caattaaaaa 2300

aaattaaaac caatttaaaa aaaaaaagaa cacaggagat tccagtctac 2350

ttgagttagc ataatacaga agtcccctct actttaactt ttacaaaaaa 2400

gtaacctgaa ctaatctgat gttaaccaat gtatttattt ctgtggttct 2450

gtttccttgt tccaatttga caaaacccac tgttcttgta ttgtattgcc 2500

cagggggagc tatcactgta cttgtagagt ggtgctgctt taattcataa 2550

atcacaaata aaagccaatt agctctaaaa aaaaaaaa 2588

<210> 10
<211> 3963
<212> DNA
<213> Homo sapiens

<400> 10
gacctggggg ggcgagggcc ccccaaactc agttcggatc ctacccgagt 50

gaggcggcgc catggagctc cgggtgctgc tctgctgggc ttcgttggcc 100

gcagctttgg aagagaccct gctgaacaca aaattggaaa ctgctgatct 150

gaagtgggtg acattccctc aggtggacgg gcagtgggag gaactgagcg 200

gcctggatga ggaacagcac agcgtgcgca cctacgaagt gtgtgacgtg 250

cagcgtgccc cgggccaggc ccactggctt cgcacaggtt gggtcccacg 300

gcggggcgcc gtccacgtgt acgccacgct gcgcttcacc atgctcgagt 350

gcctgtccct gcctcgggct gggcgctcct gcaaggagac cttcaccgtc 400

ttctactatg agagcgatgc ggacacggcc acggccctca cgccagcctg 450

gatggagaac ccctacatca aggtggacac ggtggccgcg gagcatctca 500

cccggaagcg ccctggggcc gaggccaccg ggaaggtgaa tgtcaagacg 550

ctgcgtctgg gaccgctcag caaggctggc ttctacctgg ccttccagga 600

ccagggtgcc tgcatggccc tgctatccct gcacctcttc tacaaaaagt 650

gcgcccagct gactgtgaac ctgactcgat tcccggagac tgtgcctcgg 700

gagctggttg tgcccgtggc cggtagctgc gtggtggatg ccgtccccgc 750

ccctggcccc agccccagcc tctactgccg tgaggatggc cagtgggccg 800

aacagccggt cacgggctgc agctgtgctc cggggttcga ggcagctgag 850

gggaacacca agtgccgagc ctgtgcccag ggcaccttca agcccctgtc 900

aggagaaggg tcctgccagc catgcccagc caatagccac tctaacacca 950

ttggatcagc cgtctgccag tgccgcgtcg ggtacttccg ggcacgcaca 1000

gacccccggg gtgcaccctg caccacccct ccttcggctc cgcggagcgt 1050

ggtttcccgc ctgaacggct cctccctgca cctggaatgg agtgcccccc 1100

tggagtctgg tggccgagag gacctcacct acgccctccg ctgccgggag 1150

tgccgacccg gaggctcctg tgcgccctgc gggggagacc tgacttttga 1200

ccccggcccc cgggacctgg tggagccctg ggtggtggtt cgagggctac 1250

gtcctgactt cacctatacc tttgaggtca ctgcattgaa cggggtatcc 1300

tccttagcca cggggcccgt cccatttgag cctgtcaatg tcaccactga 1350

ccgagaggta cctcctgcag tgtctgacat ccgggtgacg cggtcctcac 1400

ccagcagctt gagcctggcc tgggctgttc cccgggcacc cagtggggct 1450

gtgctggact acgaggtcaa ataccatgag aagggcgccg agggtcccag 1500

cagcgtgcgg ttcctgaaga cgtcagaaaa ccgggcagag ctgcgggggc 1550

tgaagcgggg agccagctac ctggtgcagg tacgggcgcg ctctgaggcc 1600

ggctacgggc ccttcggcca ggaacatcac agccagaccc aactggatga 1650

gagcgagggc tggcgggagc agctggccct gattgcgggc acggcagtcg 1700

tgggtgtggt cctggtcctg gtggtcattg tggtcgcagt tctctgcctc 1750

aggaagcaga gcaatgggag agaagcagaa tattcggaca aacacggaca 1800

gtatctcatc ggacatggta ctaaggtcta catcgacccc ttcacttatg 1850

aagaccctaa tgaggctgtg agggaatttg caaaagagat cgatgtctcc 1900

tacgtcaaga ttgaagaggt gattggtgca ggtgagtttg gcgaggtgtg 1950

tcgggggcgg ctcaaggccc cagggaagaa ggagagctgt gtggcaatca 2000

agaccctgaa gggtggctac acggagcggc agcggcgtga gtttctgagc 2050

gaggcctcca tcatgggcca gttcgagcac cccaatatca tccgcctgga 2100

gggcgtggtc accaacagca tgcccgtcat gattctcaca gagttcatgg 2150

agaacggcgc cctggactcc ttcctgcggc taaacgacgg acagttcaca 2200

gtcatccagc tcgtgggcat gctgcggggc atcgcctcgg gcatgcggta 2250

ccttgccgag atgagctacg tccaccgaga cctggctgct cgcaacatcc 2300

tagtcaacag caacctcgtc tgcaaagtgt ctgactttgg cctttcccga 2350

ttcctggagg agaactcttc cgatcccacc tacacgagct ccctgggagg 2400

aaagattccc atccgatgga ctgccccgga ggccattgcc ttccggaagt 2450

tcacttccgc cagtgatgcc tggagttacg ggattgtgat gtgggaggtg 2500

atgtcatttg gggagaggcc gtactgggac atgagcaatc aggacgtgat 2550

caatgccatt gaacaggact accggctgcc cccgccccca gactgtccca 2600

cctccctcca ccagctcatg ctggactgtt ggcagaaaga ccggaatgcc 2650

cggccccgct tcccccaggt ggtcagcgcc ctggacaaga tgatccggaa 2700

ccccgccagc ctcaaaatcg tggcccggga gaatggcggg gcctcacacc 2750

ctctcctgga ccagcggcag cctcactact cagcttttgg ctctgtgggc 2800

gagtggcttc gggccatcaa aatgggaaga tacgaagaaa gtttcgcagc 2850

cgctggcttt ggctccttcg agctggtcag ccagatctct gctgaggacc 2900

tgctccgaat cggagtcact ctggcgggac accagaagaa aatcttggcc 2950

agtgtccagc acatgaagtc ccaggccaag ccgggaaccc cgggtgggac 3000

aggaggaccg gccccgcagt actgacctgc aggaactccc caccccaggg 3050

acaccgcctc cccattttcc ggggcagagt ggggactcac agaggccccc 3100

agccctgtgc cccgctggat tgcactttga gcccgtgggg tgaggagttg 3150

gcaatttgga gagacaggat ttgggggttc tgccataata ggaggggaaa 3200

atcacccccc agccacctcg gggaactcca gaccaagggt gagggcgcct 3250

ttccctcagg actgggtgtg accagaggaa aaggaagtgc ccaacatctc 3300

ccagcctccc caggtgcccc cctcaccttg atgggtgcgt tcccgcagac 3350

caaagagagt gtgactccct tgccagctcc agagtggggg ggctgtccca 3400

gggggcaaga aggggtgtca gggcccagtg acaaaatcat tggggtttgt 3450

agtcccaact tgctgctgtc accaccaaac tcaatcattt ttttcccttg 3500

taaatgcccc tcccccagct gctgccttca tattgaaggt ttttgagttt 3550

tgtttttggt cttaattttt ctccccgttc ccttttttgtt tcttcgtttt 3600

gtttttctac cgtccttgtc ataactttgt gttggaggga acctgtttca 3650

ctatggcctc ctttgcccaa gttgaaacag gggcccatca tcatgtctgt 3700

ttccagaaca gtgccttggt catcccacat ccccggaccc cgcctgggac 3750

ccccaagctg tgtcctatga aggggtgtgg ggtgaggtag tgaaaagggc 3800

ggtagttggt ggtggaaccc agaaacggac gccggtgctt ggaggggttc 3850

ttaaattata tttaaaaaag taacttttg tataaataa agaaaatggg 3900

acgtgtccca aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3950

aaaaaaaaaa aga 3963

<210> 11
<211> 3476
<212> DNA
<213> Homo sapiens

<400> 11
gctctatgcc gcctaccttg ctctcgccgc tgctgccgga gccgaagcag 50

agaaggcagc gggtcccgtg accgtcccga gagccccgcg ctcccgacca 100

ggggggcggg gcggccccgg ggagggcggg gcaggggcgg ggggaagaaa 150

gggggttttg tgctgcgccg ggagggccgg cgccctcttc cgaatgtcct 200

gcggccccag cctctcctca cgctcgcgca gtctccgccg cagtctcagc 250

tgcagctgca ggactgagcc gtgcacccgg aggagacccc cggaggaggc 300

gacaaacttc gcagtgccgc gacccaaccc cagccctggg tagcctgcag 350

catggcccag ctgttcctgc ccctgctggc agccctggtc ctggcccagg 400

ctcctgcagc tttagcagat gttctggaag gagacagctc agaggaccgc 450

gcttttcgcg tgcgcatcgc gggcgacgcg ccactgcagg gcgtgctcgg 500

cggcgccctc accatccctt gccacgtcca ctacctgcgg ccaccgccga 550

gccgccgggc tgtgctgggc tctccgcggg tcaagtggac tttcctgtcc 600

cggggccggg aggcagaggt gctggtggcg cggggagtgc gcgtcaaggt 650

gaacgaggcc taccggttcc gcgtggcact gcctgcgtac ccagcgtcgc 700

tcaccgacgt ctccctggcg ctgagcgagc tgcgccccaa cgactcaggt 750

atctatcgct gtgaggtcca gcacggcatc gatgacagca gcgacgctgt 800

ggaggtcaag gtcaaagggg tcgtctttct ctaccgagag ggctctgccc 850

gctatgcttt ctccttttct ggggcccagg aggcctgtgc ccgcattgga 900

gcccacatcg ccaccccgga gcagctctat gccgcctacc ttgggggcta 950

tgagcaatgt gatgctggct ggctgtcgga tcagaccgtg aggtatccca 1000

tccagacccc acgagaggcc tgttacggag acatggatgg cttccccggg 1050

gtccggaact atggtgtggt ggacccggat gacctctatg atgtgtactg 1100

ttatgctgaa gacctaaatg gagaactgtt cctgggtgac cctccagaga 1150

agctgacatt ggaggaagca cgggcgtact gccaggagcg gggtgcagag 1200

attgccacca cgggccaact gtatgcagcc tgggatggtg gcctggacca 1250

ctgcagccca gggtggctag ctgatggcag tgtgcgctac cccatcgtca 1300

cacccagcca gcgctgtggt gggggcttgc ctggtgtcaa gactctcttc 1350

ctcttcccca accagactgg cttccccaat aagcacagcc gcttcaacgt 1400

ctactgcttc cgagactcgg cccagccttc tgccatccct gaggcctcca 1450

acccagcctc caacccagcc tctgatggac tagaggctat cgtcacagtg 1500

acagagaccc tggaggaact gcagctgcct caggaagcca cagagagtga 1550

atcccgtggg gccatctact ccatccccat catggaggac ggaggaggtg 1600

gaagctccac tccagaagac ccagcagagg cccctaggac gctcctagaa 1650

tttgaaacac aatccatggt accgcccacg gggttctcag aagaggaagg 1700

taaggcattg gaggaagaag agaaatatga agatgaagaa gagaaagagg 1750

aggaagaaga agaggaggag gtggaggatg aggctctgtg ggcatggccc 1800

agcgagctca gcagcccggg ccctgaggcc tctctcccca ctgagccagc 1850

agcccaggag aagtcactct cccaggcgcc agcaagggca gtcctgcagc 1900

ctggtgcatc accacttcct gatggagagt cagaagcttc caggcctcca 1950

agggtccatg gaccacctac tgagactctg cccactccca gggagaggaa 2000

cctagcatcc ccatcacctt ccactctggt tgaggcaaga gaggtggggg 2050

aggcaactgg tggtcctgag ctatctgggg tccctcgagg agagagcgag 2100

gagacaggaa gctccgaggg tgccccttcc ctgcttccag ccacacgggc 2150

ccctgagggt accagggagc tggaggcccc ctctgaagat aattctggaa 2200

gaactgcccc agcagggacc tcagtgcagg cccagccagt gctgcccact 2250

gacagcgcca gccgaggtgg agtggccgtg gtccccgcat caggtgactg 2300

tgtccccagc ccctgccaca atggtgggac atgcttggag gaggaggaag 2350

gggtccgctg cctatgtctg cctggctatg ggggggacct gtgcgatgtt 2400

ggcctccgct tctgcaaccc cggctgggac gccttccagg gcgcctgcta 2450

caagcacttt tccacacgaa ggagctggga ggaggcagag acccagtgcc 2500

ggatgtacgg cgcgcatctg gccagcatca gcacacccga ggaacaggac 2550

ttcatcaaca accggtaccg ggagtaccag tggatcggac tcaacgacag 2600

gaccatcgaa ggcgacttct tgtggtcgga tggcgtcccc ctgctctatg 2650

agaactggaa ccctgggcag cctgacagct acttcctgtc tggagagaac 2700

tgcgtggtca tggtgtggca tgatcaggga caatggagtg acgtgccctg 2750

caactaccac ctgtcctaca cctgcaagat ggggctggtg tcctgtgggc 2800

cgccaccgga gctgcccctg gctcaagtgt tcggccgccc acggctgcgc 2850

tatgaggtgg acactgtgct tcgctaccgg tgccgggaag gactggccca 2900

gcgcaatctg ccgctgatcc gatgccaaga gaacggtcgt tgggaggccc 2950

cccagatctc ctgtgtgccc agaagacctg cccgagctct gcacccagag 3000

gaggacccag aaggacgtca ggggaggcta ctgggacgct ggaaggcgct 3050

gttgatcccc ccttccagcc ccatgccagg tccctagggg gcaaggcctt 3100

gaacactgcc ggccacagca ctgccctgtc acccaaattt tccctcacac 3150

cttgcgctcc cgccaccaca ggaagtgaca acatgacgag gggtggtgct 3200

ggagtccagg tgacagttcc tgaaggggct tctgggaaat acctaggagg 3250

ctccagccca gcccaggccc tctcccccta ccctgggcac cagatcttcc 3300

atcagggccg gagtaaatcc ctaagtgcct caactgccct ctccctggca 3350

gccatcttgt cccctctatt cctctaggga gcactgtgcc cactctttct 3400

gggttttcca agggaatggg cttgcaggat ggagtgtctg taaaatcaac 3450

aggaaataaa actgtgtatg agccca 3476

<210> 12
<211> 2558
<212> DNA
<213> Homo sapiens

<400> 12
tgtggcactg cctgcgtacc caaccccagc cctgggtagc ctgcagcatg 50

gcccagctgt tcctgcccct gctggcagcc ctggtcctgg cccaggctcc 100

tgcagcttta gcagatgttc tggaaggaga cagctcagag gaccgcgctt 150

ttcgcgtgcg catcgcgggc gacgcgccac tgcagggcgt gctcggcggc 200

gccctcacca tcccttgcca cgtccactac ctgcggccac cgccgagccg 250

ccgggctgtg ctgggctctc cgcgggtcaa gtggactttc ctgtcccggg 300

gccgggaggc agaggtgctg gtggcgcggg gagtgcgcgt caaggtgaac 350

gaggcctacc ggttccgcgt ggcactgcct gcgtacccag cgtcgctcac 400

cgacgtctcc ctggcgctga gcgagctgcg ccccaacgac tcaggtatct 450

atcgctgtga ggtccagcac ggcatcgatg acagcagcga cgctgtggag 500

gtcaaggtca aaggggtcgt ctttctctac cgagagggct ctgcccgcta 550

tgctttctcc ttttctgggg cccaggaggc ctgtgcccgc attggagccc 600

acatcgccac cccggagcag ctctatgccg cctaccttgg gggctatgag 650

caatgtgatg ctggctggct gtcggatcag accgtgaggt atcccatcca 700

gaccccacga gaggcctgtt acggagacat ggatggcttc cccggggtcc 750

ggaactatgg tgtggtggac ccggatgacc tctatgatgt gtactgttat 800

gctgaagacc taaatggaga attgttcctg ggtgaccctc cagagaagct 850

gacattggag gaagcacggg cgtactgcca ggagcggggt gcagagattg 900

ccaccacggg ccaactgtat gcagcctggg atggtggcct ggaccactgc 950

agcccagggt ggctagctga tggcagtgtg cgctacccca tcgtcacacc 1000

cagccagcgc tgtggtgggg gcttgcctgg tgtcaagact ctcttcctct 1050

tccccaacca gactggcttc cccaataagc acagccgctt caacgtctac 1100

tgcttccgag actcggccca gccttctgcc atccctgagg cctccaaccc 1150

agcctccaac ccagcctctg atggactaga ggctatcgtc acagtgacag 1200

agaccctgga ggaactgcag ctgcctcagg aagccacaga gagtgaatcc 1250

cgtggggcca tctactccat ccccatcatg gaggacggag gaggtggaag 1300

ctccactcca gaagacccag cagaggcccc taggacgctc ctagaatttg 1350

aaacacaatc catggtaccg cccacggggt tctcagaaga ggaaggtaag 1400

gcattggagg aagaagagaa atatgaagat gaagaagaga aagaggagga 1450

agaagaagag gaggaggtgg aggatgaggc tctgtgggca tggcccagcg 1500

agctcagcag cccgggccct gaggcctctc tccccactga gccagcagcc 1550

caggaggagt cactctccca ggcgccagca agggcagtcc tgcagcctgg 1600

tgcatcacca cttcctgatg gagagtcaga agcttccagg cctccaaggg 1650

tccatggacc acctactgag actctgccca ctcccaggga gaggaaccta 1700

gcatccccat caccttccac tctggttgag gcaagagagg tgggggaggc 1750

aactggtggt cctgagctat ctggggtccc tcgaggagag agcgaggaga 1800

caggaagctc cgagggtgcc ccttccctgc ttccagccac acgggcccct 1850

gagggtacca gggagctgga ggccccctct gaagataatt ctggaagaac 1900

tgccccagca gggacctcag tgcaggccca gccagtgctg cccactgaca 1950

gcgccagccg aggtggagtg gccgtggtcc ccgcatcagg taattctgcc 2000

caaggctcaa ctgccctctc tatcctactc cttttcttcc ccctgcagct 2050

ctgggtcacc tgacctgtag tcctttaacc caccatcatc ccaaactctc 2100

ctgtcctttg ccttcattct cttacccacc tctacctatg ggtctccaat 2150

ctcggatatc caccttgtgg gtatctcagc tctccgcgtc tttaccctgt 2200

gatcccagcc ccgccactga ccatctgtga cccttccctg ccattgggcc 2250

ctccacctgt ggctcacatc tcgccagccc cacagagcat cctcaggcct 2300

ctccaagggt cctcatcacc tattgcagcc ttcagggctc ggcctatttt 2350

ccactactcc cttcatccgc ctgtgtgccg tccccttttag ctgcctccta 2400

ttgatctcag ggaagcctgg gagtcccttc tcacccctca acctccggag 2450

tccaggagaa cccgtacccc cacagagcct taagcaacta cttctgtgaa 2500

gtatttttg actgtttcat ggaaaacaag ccttggaaat aaatctctat 2550

taaaccgc 2558

<210> 13
<211> 401
<212> DNA
<213> Homo sapiens

<400> 13
agccagctca ggctacacta tcccaggatc agcatggccg tccgccagtg 50

ggtaatcgcc ctggccttgg ctgccctcct tgttgtggac agggaagtgc 100

cagtggcagc aggaaagctc cctttctcaa gaatgcccat ctgtgaacac 150

atggtagagt ctccaacctg ttcccagatg tccaacctgg tctgcggcac 200

tgatgggctc acatatacga atgaatgcca gctctgcttg gcccggataa 250

aaaccaaaca ggacatccag atcatgaaag atggcaaatg ctgatcccac 300

aggagcacct caagccatga agtgtcagct ggagaacagt ggtgggcatg 350

gagaggatat gacatgaaat aaaagatcca gcccaaaaaa aaaagaaaaa 400

a 401

<210> 14
<211> 1572
<212> DNA
<213> Homo sapiens

<400> 14
cgtttccaat gcacgtacag cccgtacaca ccgtgtgctg ggacacccca 50

cagtcagccg catggctccc ctgtgcccca gcccctggct ccctctgttg 100

atcccggccc ctgctccagg cctcactgtg caactgctgc tgtcactgct 150

gcttctggtg cctgtccatc cccagaggtt gccccggatg caggaggatt 200

ccccccctggg aggaggctct tctggggaag atgacccact gggcgaggag 250

gatctgccca gtgaagagga ttcacccaga gaggaggatc cacccggaga 300

ggaggatcta cctggagagg aggatctacc tggagaggag gatctacctg 350

aagttaagcc taaatcagaa gaagagggct ccctgaagtt agaggatcta 400

cctactgttg aggctcctgg agatcctcaa gaaccccaga ataatgccca 450

cagggacaaa gaaggggatg accagagtca ttggcgctat ggaggcgacc 500

cgccctggcc ccgggtgtcc ccagcctgcg cgggccgctt ccagtccccg 550

gtggatatcc gcccccagct cgccgccttc tgcccggccc tgcgcccccc 600

ggaactcctg ggcttccagc tcccgccgct cccagaactg cgcctgcgca 650

acaatggcca cagtgtgcaa ctgaccctgc ctcctgggct agagatggct 700

ctgggtcccg ggcgggagta ccgggctctg cagctgcatc tgcactgggg 750

ggctgcaggt cgtccgggct cggagcacac tgtggaaggc caccgtttcc 800

ctgccgagat ccacgtggtt cacctcagca ccgcctttgc cagagttgac 850

gaggccttgg ggcgcccggg aggcctggcc gtgttggccg cctttctgga 900

ggagggcccg gaagaaaaca gtgcctatga gcagttgctg tctcgcttgg 950

aagaaatcgc tgaggaaggc tcagagactc aggtcccagg actggacata 1000

tctgcactcc tgccctctga cttcagccgc tacttccgat atgaggggtc 1050

tctgactaca ccgccctgtg cccagggtgt catctggact gtgtttaacc 1100

agacagtgat gctgagtgct aagcagctcc acaccctctc tgacaccctg 1150

tggggacctg gtgactctcg gctacagctg aacttccgag cgacgcagcc 1200

tttgaatggg cgagtgattg aggcctcctt ccctgctgga gtggacagca 1250

gtcctcgggc tgctgagcca gtccagctga attcctgcct ggctgctggt 1300

gacatcctag ccctggtttt tggcctcctt tttgctgtca ccagcgtcgc 1350

gttccttgtg cagatgagaa ggcagcacag aaggggaacc aaaggggtg 1400

tgagctaccg cccagcagag gtagccgaga ctggagccta gaggctggat 1450

cttggagaat gtgagaagcc agccagaggc atctgagggg gagccggtaa 1500

ctgtcctgtc ctgctcatta tgccacttcc ttttaactgc caagaaattt 1550

tttaaaataa atatttataa ta 1572

<210> 15
<211> 2573
<212> DNA
<213> Homo sapiens

<400> 15
tccacatcca ggtctttgtg ctcctcgctt gcctgttcct tttccacgca 50

ttttccagga taactgtgac tccaggcccg caatggatgc cctgcaacta 100

gcaaattcgg cttttgccgt tgatctgttc aaacaactat gtgaaaagga 150

gccactgggc aatgtcctct tctctccaat ctgtctctcc acctctctgt 200

cacttgctca agtgggtgct aaaggtgaca ctgcaaatga aattggacag 250

gttcttcatt ttgaaaatgt caaagatgta ccctttggat ttcaaacagt 300

aacatcggat gtaaacaaac ttagttcctt ttactcactg aaactaatca 350

agcggctcta cgtagacaaa tctctgaatc tttctacaga gttcatcagc 400

tctacgaaga gaccgtatgc aaaggaattg gaaactgttg acttcaaaga 450

taaattggaa gaaacgaaag gtcagatcaa caactcaatt aaggatctca 500

cagatggcca ctttgagaac attttagctg acaacagtgt gaacgaccag 550

accaaaatcc ttgtggttaa tgctgcctac tttgttggca agtggatgaa 600

gaaattttct gaatcagaaa caaaagaatg tcctttcaga gtcaacaaga 650

cagacaccaa accagtgcag atgatgaaca tggaggccac gttctgtatg 700

ggaaacattg acagtatcaa ttgtaagatc atagagcttc cttttcaaaa 750

taagcatctc agcatgttca tcctactacc caaggatgtg gaggatgagt 800

ccacaggctt ggagaagatt gaaaaacaac tcaactcaga gtcactgtca 850

cagtggacta atcccagcac catggccaat gccaaggtca aactctccat 900

tccaaaattt aaggtggaaa agatgattga tcccaaggct tgtctggaaa 950

atctagggct gaaacatatc ttcagtgaag acacatctga tttctctgga 1000

atgtcagaga ccaagggagt ggccctatca aatgttatcc acaaagtgtg 1050

cttagaaata actgaagatg gtggggattc catagaggtg ccaggagcac 1100

ggatcctgca gcacaaggat gaattgaatg ctgaccatcc ctttatttac 1150

atcatcaggc acaacaaaac tcgaaacatc attttctttg gcaaattctg 1200

ttctccttaa gtggcatagc ccatgttaag tcctccctga cttttctgtg 1250

gatgccgatt tctgtaaact ctgcatccag agattcattt tctagataca 1300

ataaattgct aatgttgctg gatcaggaag ccgccagtac ttgtcatatg 1350

tagccttcac acagatagac ctttttttt tttccaattc tatcttttgt 1400

ttccttttt cccataagac aatgacatac gcttttaatg aaaaggaatc 1450

acgttagagg aaaaatattt attcattatt tgtcaaattg tccgggttag 1500

ttggcagaaa tacagtcttc cacaaagaaa attcctataa ggaagatttg 1550

gaagctcttc ttcccagcac tatgctttcc ttctttggga tagagaatgt 1600

tccagacatt ctcgcttccc tgaaagactg aagaaagtgt agtgcatggg 1650

acccacgaaa ctgccctggc tccagtgaaa cttgggcaca tgctcaggct 1700

actataggtc cagaagtcct tatgttaagc cctggcaggc aggtgtttat 1750

taaaattctg aattttgggg attttcaaaa gataatattt tacatacact 1800

gtatgttata gaacttcatg gatcagatct ggggcagcac cctataaatc 1850

aacaccttaa tatgctgcaa caaaatgtag aatattcaga caaaatggat 1900

acataaagac taagtagccc ataaggggtc aaaatttgct gccaaatgcg 1950

tatgccacca acttacaaaa acacttcgtt cgcagagctt ttcagattgt 2000

ggaatgttgg ataaggaatt atagacctct agtagctgaa atgcaagacc 2050

ccaagaggaa gttcagatct taatataaat tcactttcat ttttgatagc 2100

tgtcccatct ggtcatttgg ttggcactag actggtggca ggggcttcta 2150

gctgacttgc acagggattc tcacaatagc cgatatcaga atttgtgttg 2200

aaggaacttg tctcttcatc taatatgata gcgggaaaag gagaggaaac 2250

tactgccttt agaaaatata agtaaagtga ttaaagtgct cacgttacct 2300

tgacacatag tttttcagtc tatgggttta gttactttag atggcaagca 2350

tgtaacttat attaatagta atttgtaaag ttggttggat aagctatccg 2400

tgttgcaggt tcatggatta cttctctata aaaaatatgt atttaccaaa 2450

aattttgtga cattccttct cccatctctt ccttgacctg cattgtaaat 2500

aggttcttct tgttctgaga ttcaatattg aatttttcct atgctattga 2550

caataaaata ttattgaact aca 2573

<210> 16
<211> 1145
<212> DNA
<213> Homo sapiens

<400> 16
gccgagccgg ggctccggaa gccggcgggg gcgccgcggc cgtgcggggc 50

gtcaatggat cgccactcca gctacatctt catctggctg cagctggagc 100

tctgcgccat ggccgtgctg ctcaccaaag gtgaaattcg atgctactgt 150

gatgctgccc actgtgtagc cactggttat atgtgtaaat ctgagctcag 200

cgcctgcttc tctagacttc ttgatcctca gaactcaaat tccccactca 250

cccatggctg cctggactct cttgcaagca cgacagacat ctgccaagcc 300

aaacaggccc gaaaccactc tggcaccacc atacccacat tggaatgctg 350

tcatgaagac atgtgcaatt acagagggct gcacgatgtt ctctctcctc 400

ccaggggtga ggcctcagga caaggaaaca ggtatcagca tgatggtagc 450

agaaacctta tcaccaaggt gcaggagctg acttcttcca aagagttgtg 500

gttccgggca gcggtcattg ccgtgcccat tgctggaggg ctgattttag 550

tgttgcttat tatgttggcc ctgaggatgc ttcgaagtga aaataagagg 600

ctgcaggatc agcggcaaca gatgctctcc cgtttgcact acagctttca 650

cggacaccat tccaaaaagg ggcaggttgc aaagttagac ttggaatgca 700

tggtgccggt cagtgggcac gagaactgct gtctgacctg tgataaaatg 750

agacaagcag acctcagcaa cgataagatc ctctcgcttg ttcactgggg 800

catgtacagt gggcacggga agctggaatt cgtatgacgg agtcttatct 850

gaactacact tactgaacag cttgaaggcc ttttgagttc tgctggacag 900

gagcacttta tctgaagaca aactcattta atcatctttg acagacaaaa 950

tgacctctgc aaacagaatc ttggatattt cttctgaagg attatttgca 1000

cagacttaaa tacagttaaa tgtgttattt gcttttaaaa ttataaaaag 1050

caaagagaag actttgtaca cactgtcacc agggttattt gcatccaagg 1100

gagctggaat tgagtaccta aataaacaaa aatgtgccct aaaaa 1145

<210> 17
<211> 5668
<212> DNA
<213> Homo sapiens

<400> 17
gtcacttagg aaaaggtgtc ctttcgggca gccgggctca gcatgaggaa 50

cagaaggaat gacactctgg acagcacccg gaccctgtac tccagcgcgt 100

ctcggagcac agacttgtct tacagtgaaa gcgacttggt gaattttatt 150

caagcaaatt ttaagaaacg agaatgtgtc ttctttacca aagattccaa 200

ggccacggag aatgtgtgca agtgtggcta tgcccagagc cagcacatgg 250

aaggcaccca gatcaaccaa agtgagaaat ggaactacaa gaaacacacc 300

aaggaatttc ctaccgacgc ctttggggat attcagtttg agacactggg 350

gaagaaaggg aagtatatac gtctgtcctg cgacacggac gcggaaatcc 400

tttacgagct gctgacccag cactggcacc tgaaaacacc caacctggtc 450

atttctgtga ccgggggcgc caagaacttc gccctgaagc cgcgcatgcg 500

caagatcttc agccggctca tctacatcgc gcagtccaaa ggtgcttgga 550

ttctcacggg aggcacccat tatggcctga cgaagtacat cggggaggtg 600

gtgagagata acaccatcag caggagttca gaggagaata ttgtggccat 650

tggcatagca gcttggggca tggtctccaa ccgggacacc ctcatcagga 700

attgcgatgc tgagggctat tttttagccc agtaccttat ggatgacttc 750

acaagggatc cactgtatat cctggacaac aaccacacac atttgctgct 800

cgtggacaat ggctgtcatg gacatcccac tgtcgaagca aagctccgga 850

atcagctaga gaagcatatc tctgagcgca ctattcaaga ttccaactat 900

ggtggcaaga tccccattgt gtgttttgcc caaggaggtg gaaaagagac 950

tttgaaagcc atcaatacct ccatcaaaaa taaaattcct tgtgtggtgg 1000

tggaaggctc gggccggatc gctgatgtga tcgctagcct ggtggaggtg 1050

gaggatgccc cgacatcttc tgccgtcaag gagaagctgg tgcgcttttt 1100

accccgcacg gtgtcccggc tgtctgagga ggagactgag agttggatca 1150

aatggctcaa agaaattctc gaatgttctc acctattaac agttattaaa 1200

atggaagaag ctggggatga attgtgagc aatgccatct cctacgctct 1250

atacaaagcc ttcagcacca gtgagcaaga caaggataac tggaatgggc 1300

agctgaagct tctgctggag tggaaccagc tggacttagc caatgatgag 1350

attttcacca atgaccgccg atgggagtct gctgaccttc aagaagtcat 1400

gtttacggct ctcataaagg acagacccaa gtttgtccgc ctctttctgg 1450

agaatggctt gaacctacgg aagtttctca cccatgatgt cctcactgaa 1500

ctcttctcca accacttcag cacgcttgtg taccggaatc tgcagatcgc 1550

caagaattcc tataatgatg ccctcctcac gtttgtctgg aaactggttg 1600

cgaacttccg aagaggcttc cggaaggaag acagaaatgg ccgggacgag 1650

atggacatag aactccacga cgtgtctcct attactcggc acccctgca 1700

agctctcttc atctgggcca ttcttcagaa taagaaggaa ctctccaaag 1750

tcatttggga gcagaccagg ggctgcactc tggcagccct gggagccagc 1800

aagcttctga agactctggc caaagtgaag aacgacatca atgctgctgg 1850

ggagtccgag gagctggcta atgagtacga gacccgggct gttgagctgt 1900

tcactgagtg ttacagcagc gatgaagact tggcagaaca gctgctggtc 1950

tattcctgtg aagcttgggg tggaagcaac tgtctggagc tggcggtgga 2000

ggccacagac cagcatttca ccgcccagcc tggggtccag aattttcttt 2050

ctaagcaatg gtatggagag atttcccgag acaccaagaa ctggaagatt 2100

atcctgtgtc tgtttattat acccttggtg ggctgtggct ttgtatcatt 2150

taggaagaaa cctgtcgaca agcacaagaa gctgctttgg tactatgtgg 2200

cgttcttcac ctccccccttc gtggtcttct cctggaatgt ggtcttctac 2250

atcgccttcc tcctgctgtt tgcctacgtg ctgctcatgg atttccattc 2300

ggtgccacac ccccccgagc tggtcctgta ctcgctggtc tttgtcctct 2350

tctgtgatga agtgagacag tggtacgtaa atggggtgaa ttattttact 2400

gacctgtgga atgtgatgga cacgctgggg ctttttttact tcatagcagg 2450

aattgtattt cggctccact cttctaataa aagctctttg tattctggac 2500

gagtcatttt ctgtctggac tacattattt tcactctaag attgatccac 2550

attttactg taagcagaaa cttaggaccc aagattataa tgctgcagag 2600

gatgctgatc gatgtgttct tcttcctgtt cctctttgcg gtgtggatgg 2650

tggcctttgg cgtggccagg caagggatcc ttaggcagaa tgagcagcgc 2700

tggaggtgga tattccgttc ggtcatctac gagccctacc tggccatgtt 2750

cggccaggtg cccagtgacg tggatggtac cacgtatgac tttgcccact 2800

gcaccttcac tgggaatgag tccaagccac tgtgtgtgga gctggatgag 2850

cacaacctgc cccggttccc cgagtggatc accatccccc tggtgtgcat 2900

ctacatgtta tccaccaaca tcctgctggt caacctgctg gtcgccatgt 2950

ttggctacac ggtgggcacc gtccaggaga acaatgacca ggtctggaag 3000

ttccagaggt acttcctggt gcaggagtac tgcagccgcc tcaatatccc 3050

cttccccttc atcgtcttcg cttacttcta catggtggtg aagaagtgct 3100

tcaagtgttg ctgcaaggag aaaaacatgg agtcttctgt ctgctgtttc 3150

aaaaatgaag acaatgagac tctggcatgg gagggtgtca tgaaggaaaa 3200

ctaccttgtc aagatcaaca caaaagccaa cgacacctca gaggaaatga 3250

ggcatcgatt tagacaactg gatacaaagc ttaatgatct caagggtctt 3300

ctgaaagaga ttgctaataa aatcaaataa aactgtatga aactctaatg 3350

gagaaaaatc taattatagc aagatcatat taaggaatgc tgatgaacaa 3400

ttttgctatc gactactaaa tgagagattt tcagacccct gggtacatgg 3450

tggatgattt taaatcaccc tagtgtgctg agaccttgag aataaagtgt 3500

gtgattggtt tcatacttga agacggatat aaaggaagaa tatttccttt 3550

atgtgtttct ccagaatggt gcctgtttct ctctgtgtct caatgcctgg 3600

gactggaggt tgatagttta agtgtgttct taccgcctcc tttttccttt 3650

aatcttattt ttgatgaaca catatatagg agaacatcta tcctatgaat 3700

aagaacctgg tcatgcttta ctcctgtatt gttattttgt tcatttccaa 3750

ttgattctct acttttccct tttttgtatt atgtgactaa ttagttggca 3800

tattgttaaa agtctctcaa attaggccag attctaaaac atgctgcagc 3850

aagaggaccc cgctctcttc aggaaaagtg ttttcatttc tcaggatgct 3900

tcttacctgt cagaggaggt gacaaggcag tctcttgctc tcttggactc 3950

accaggctcc tattgaagga accaccccca ttcctaaata tgtgaaaagt 4000

cgcccaaaat gcaaccttga aaggcactac tgactttgtt cttattggat 4050

actcctctta tttattattt ttccattaaa aataatagct ggctattata 4100

gaaaatttag accatacaga gatgtagaaa gaacataaat tgtccccatt 4150

accttaaggt aatcactgct aacaatttct ggatggtttt tcaagtctat 4200

ttttttttcta tgtatgtctc aattctcttt caaaatttta cagaatgtta 4250

tcatactaca tatatacttt ttatgtaagc tttttcactt agtattttat 4300

caaatatgtt tttattatat tcatagcctt cttaaacatt atatcaataa 4350

ttgcataata ggcaacctct agcgattacc ataattttgc tcattgaagg 4400

ctatctccag ttgatcattg ggatgagcat ctttgtgcat gaatcctatt 4450

gctgtatttg ggaaaatttt ccaaggttag attccaataa atatctattt 4500

attattaaat attaaaatat cgatttatta ttaaaaccat ttataaggct 4550

ttttcataaa tgtatagcaa ataggaatta ttaacttgag cataagatat 4600

gagatacatg aacctgaact attaaaataa aatattatat ttaaccctag 4650

tttaagaaga agtcaatatg cttatttaaa tattatggat ggtgggcaga 4700

tcacttgagg tcaggagttc gagaccagcc tggccaacat ggcaaaacca 4750

catctctact aaaaataaaa aaattagctg ggtgtggtgg tgcactcctg 4800

taatcccagc tactcagaag gctgaggtac aagaattgct ggaacctggg 4850

aggcggaggt tgcagtgaac caagattgca ccactgcact ccagccgggg 4900

tgacagagtg agactccgac tgaaaataaa taaataaata aataaataaa 4950

taaataaata aatattatgg atggtgaagg gaatggtata gaattggaga 5000

gattatctta ctgaacacct gtagtcccag ctttctctgg aagtggtggt 5050

atttgagcag gatgtgcaca aggcaattga aatgcccata attagtttct 5100

cagctttgaa tacactataa actcagtggc tgaaggagga aattttagaa 5150

ggaagctact aaaagatcta atttgaaaaa ctacaaaagc attaactaaa 5200

aaagtttatt ttccttttgt ctgggcagta gtgaaaataa ctactcacaa 5250

cattcactat gtttgcaagg aattaacaca aataaaagat gccttttttac 5300

ttaaacgcca agacagaaaa cttgcccaat actgagaagc aacttgcatt 5350

agagagggaa ctgttaaatg ttttcaaccc agttcatctg gtggatgttt 5400

ttgcaggtta ctctgagaat tttgcttatg aaaaatcatt atttttagtg 5450

tagttcacaa taatgtattg aacatacttc taatcaaagg tgctatgtcc 5500

ttgtgtatgg tactaaatgt gtcctgtgta cttttgcaca actgagaatc 5550

ctgcggcttg gtttaatgag tgtgttcatg aaataaataa tggaggaatt 5600

gtcaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5650

aaaaaaaaaa aaaaaaaa 5668

<210> 18
<211> 3740
<212> DNA
<213> Homo sapiens

<400> 18
gaattcgccc ttggctcagc atgaggaaca gaaggaatga cactctggac 50

agcacccgga ccctgtactc cagcgcgtct cggagcacag acttgtctta 100

cagtgaaagc gacttggtga attttattca agcaaatttt aagaaacgag 150

aatgtgtctt ctttaccaaa gattccaagg ccacggagaa tgtgtgcaag 200

tgtggctatg cccagagcca gcacatggaa ggcacccaga tcaaccaaag 250

tgagaaatgg aactacaaga aacacaccaa ggaatttcct accgacgcct 300

ttggggatat tcagtttgag acactgggga agaaagggaa gtatatacgt 350

ctgtcctgcg acacggacgc ggaaatcctt tacgagctgc tgacccagca 400

ctggcacctg aaaacaccca acctggtcat ttctgtgacc gggggcgcca 450

agaacttcgc cctgaagccg cgcatgcgca agatcttcag ccggctcatc 500

tacatcgcgc agtccaaagg tgcttggatt ctcacgggag gcacccatta 550

tggcctgatg aagtacatcg gggaggtggt gagagataac accatcagca 600

ggagttcaga ggagaatatt gtggccattg gcatagcagc ttggggcatg 650

gtctccaacc gggacaccct catcaggaat tgcgatgctg agggctattt 700

tttagcccag taccttatgg atgacttcac aagagatcca ctgtatatcc 750

tggacaacaa ccacacacat ttgctgctcg tggacaatgg ctgtcatgga 800

catcccactg tcgaagcaaa gctccggaat cagctagaga agtatatctc 850

tgagcgcact attcaagatt ccaactatgg tggcaagatc cccattgtgt 900

gttttgccca aggaggtgga aaagagactt tgaaagccat caatacctcc 950

atcaaaaata aaattccttg tgtggtggtg gaaggctcgg gccagatcgc 1000

tgatgtgatc gctagcctgg tggaggtgga ggatgccctg acatcttctg 1050

ccgtcaagga gaagctggtg cgctttttac cccgcacggt gtcccggctg 1100

cctgaggagg agactgagag ttggatcaaa tggctcaaag aaattctcga 1150

atgttctcac ctattaacag ttattaaaat ggaagaagct ggggatgaaa 1200

ttgtgagcaa tgccatctcc tacgctctat acaaagcctt cagcaccagt 1250

gagcaagaca aggataactg gaatgggcag ctgaagcttc tgctggagtg 1300

gaaccagctg gacttagcca atgatgagat tttcaccaat gaccgccgat 1350

gggagtctgc tgaccttcaa gaagtcatgt ttacggctct cataaaggac 1400

agacccaagt ttgtccgcct ctttctggag aatggcttga acctacggaa 1450

gtttctcacc catgatgtcc tcactgaact cttctccaac cacttcagca 1500

cgcttgtgta ccggaatctg cagatcgcca agaattccta taatgatgcc 1550

ctcctcacgt ttgtctggaa actggttgcg aacttccgaa gaggcttccg 1600

gaaggaagac agaaatggcc gggacgagat ggacatagaa ctccacgacg 1650

tgtctcctat tactcggcac cccctgcaag ctctcttcat ctgggccatt 1700

cttcagaata agaaggaact ctccaaagtc atttgggagc agaccagggg 1750

ctgcactctg gcagccctgg gagccagcaa gcttctgaag actctggcca 1800

aagtgaagaa cgacatcaat gctgctgggg agtccgagga gctggctaat 1850

gagtacgaga cccgggctgt tgagctgttc actgagtgtt acagcagcga 1900

tgaagacttg gcagaacagc tgctggtcta ttcctgtgaa gcttggggtg 1950

gaagcaactg tctggagctg gcggtggagg ccacagacca gcatttcatc 2000

gcccagcctg gggtccagaa ttttctttct aagcaatggt atggagagat 2050

ttcccgagac accaagaact ggaagattat cctgtgtctg tttattatac 2100

ccttggtggg ctgtggcttt gtatcattta ggaagaaacc tgtcgacaag 2150

cacaagaagc tgctttggta ctatgtggcg ttcttcacct ccccttcgt 2200

ggtcttctcc tggaatgtgg tcttctacat cgccttcctc ctgctgtttg 2250

cctacgtgct gctcatggat ttccattcgg tgccacaccc ccccgagctg 2300

gtcctgtact cgctggtctt tgtcctcttc tgtgatgaag tgagacagtg 2350

gtacgtaaat ggggtgaatt attttactga cctgtggaat gtgatggaca 2400

cgctggggct tttttacttc atagcaggaa ttgtatttcg gctccactct 2450

tctaataaaa gctctttgta ttctggacga gtcattttct gtctggacta 2500

cattattttc actctaagat tgatccacat ttttactgta agcagaaact 2550

taggacccaa gattataatg ctgcagagga tgctgatcga tgtgttcttc 2600

ttcctgttcc tctttgcggt gtggatggtg gcctttggcg tggccaggca 2650

agggatcctt aggcagaatg agcagcgctg gaggtggata ttccgttcgg 2700

tcatctacga gccctacctg gccatgttcg gccaggtgcc cagtgacgtg 2750

gatggtacca cgtatgactt tgcccactgc accttcactg ggaatgagtc 2800

caagccactg tgtgtggagc tggatgagca caacctgccc cggttccccg 2850

agtggatcac catccccctg gtgtgcatct acatgttatc caccaacatc 2900

ctgctggtca acctgctggt cgccatgttt ggctacacgg tgggcaccgt 2950

ccaggagaac aatgaccagg tctggaagtt ccagaggtac ttcctggtgc 3000

aggagtactg cagccgcctc aatatcccct tccccttcat tgtcttcgct 3050

tacttctaca tggtggtgaa gaagtgcttc aagtgttgct gcaaggagaa 3100

aaacatggag tcttctgtct gctgtttcaa aaatgaagac aatgagactc 3150

tggcatggga gggtgtcatg aaggaaaact accttgtcaa gatcaacaca 3200

aaagccaacg acacctcaga ggaaatgagg catcgattta gacaactgga 3250

tacaaagctt aatgatctca agggtcttct gaaagagatt gctaataaaa 3300

tcaaataaaa ctgtatgaac tctaatggag aaaaatctaa ttatagcaag 3350

atcatattaa ggaatgctga tgaacaattt tgctatcgac tactaaatga 3400

gagattttca gacccctggg tacatggtgg atgattttaa atcaccctag 3450

tgtgctgaga ccttgagaat aaagtgtgtg attggtttca tacttgaaga 3500

cggatataaa ggaagaatat ttcctttatg tgtttctcca gaatggtgcc 3550

tgtttctctc tgtgtctcaa tgcctgggac tggaggttga tagtttaagt 3600

gtgttcttac cgcctccttt ttcctttaat cttatttttg atgaacacat 3650

atataggaga acatctatcc tatgaataag aacctggtca tgctttaaaa 3700

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3740

<210> 19
<211> 3809
<212> DNA
<213> Homo sapiens

<400> 19
gcggcggggg cgaggtgagg tgttggcagt ggaaaggggt tcgggctcgg 50

ggggcggggg gacgcggagc ggatggcccg cgccggccgc aggggcggat 100

aaaaagccgt cgcgctgcgg gagtgggcgg gagggagagg gggtgtccga 150

gggccacaag agtatgacgg ggctgtacga gctggtgtgg cgggtgctgc 200

acgcgctgct ctgtctgcac cgcacgctca cctcctggct ccgcgttcgg 250

ttcggcacct ggaactggat ctggcggcgc tgctgccgcg ccgcctctgc 300

cgcggtccta gcgccgctcg gcttcacgct ccgcaagccc ccggcagtcg 350

gcaggaaccg ccgtcaccac cggcacccgc gcgggggggtc gtgcctggca 400

gccgcacacc accggatgcg ctggcgcgcg gacggtcgtt ccttggagaa 450

gctgcctgtg catatgggcc tggtgatcac cgaggtggag caggaaccca 500

gcttctcgga catcgcgagc ctcgtggtgt ggtgtatggc cgtgggcatc 550

tcctacatta gcgtctacga ccaccaaggt attttcaaaa gaaataattc 600

cagattgatg gatgaaattt taaaacaaca gcaagaactt ctgggcctag 650

attgttcaaa atactcacca gaatttgcaa atagtaatga caaagatgat 700

caagttttaa attgccattt ggcagtgaag gtgctgtctc cggaagatgg 750

aaaagcagat attgtaagag ctgctcagga cttttgccag ttagtagccc 800

agaagcaaag gagacccaca gatttggatg tagatacgtt agccagttta 850

cttagttcaa atggttgtcc tgatcctgat ttagtattga agttcggtcc 900

tgtggacagc acattaggct ttcttccctg gcacatcaga ttgactgaga 950

ttgtctcttt gccttcccac ctaaacatca gttatgagga ctttttctct 1000

gcccttcgtc aatatgcagc ctgtgaacag cgtctgggaa agtagtggtc 1050

attggttgca taatttgatt tgaggcttgt ggaggaaagg aaccaagtga 1100

ctctgatgtt tacaaagcac ctatgaaacc ctgtacacac ctatgaaacc 1150

ctgtacacac ctagttcata atcctcataa tttatcaaca aacacaaaga 1200

agtgtcttac ttgagagtga gtgtgtgcgt gtgtgcgtgc acacatgtgc 1250

acgtttgtat gtgtggaaat aaacataaat ggggacgtgt tggagaagga 1300

aatacataga cctacaactt tgagcatata gcagtgatgt tttaggaact 1350

gaaatgtcac acttaataaa gtcttcagcc cagctacttc cctgttttcg 1400

tggggagaag ggggcctgat tagaactgtt ctggttgtgt ttggcgggag 1450

gggaataaat tttgttcagt ccttcttagt gaccaaactt taatttttag 1500

gaataacata ttgacttact gaactgaagc attctgagtt gaaaggagcc 1550

ccagaggaaa ggagttctgt gttgctcaca tgttaaaagc ttgctcacct 1600

tcagagcaga gggaatacct atcttcagat atccgcccat tttcatctct 1650

tcattatagt caaacagtgt gacttgagag tgttgctctg gtgtctgtat 1700

tctggcttat gaagattatt tgaaaaagaa ctcttactac attgaaatgc 1750

agacttttaa aaatttaaat attggattag gcagtcaaaa aaccaaacaa 1800

gcataaaagg tcagtaagtt gtaatcttaa aagtaaaggt ggaaaactca 1850

ttataaatgg aagaaaagtt ttgatttcct tttttgtttg atgggcagta 1900

tgccatatta tatccaaagt tggtttaaaa aatacttcca tcaactattt 1950

ttatttaaaa taaacatttg agggaagtta ccaaggcagc ttttttcctc 2000

aaaagtaacc tgttcctctt tggaacagca cattttaggg gcatggttaa 2050

tacctgagat ttttactcag taaacctgat ggttaccgtg tgtaaaatat 2100

ctttaagtag gattgaaggc ctctgtgggg aatataatgt taccaaagcc 2150

tataaaaata aattttacat gttctttttt atgacagaga gcagcactgg 2200

ttctgttatt ttttaaaatg aataactgat ttcttgacag gtgtttagta 2250

tttcttccct cactgctgat tcttggatag aaaccattct ttatatttga 2300

tagactgttt tcagaaaact cttatcaaca agtgtacaat agttatctaa 2350

aactgtacat ttagaatgga gcagtttaat actagatctc agaaatttga 2400

aaaatagcaa agaagactgg atttggaaag catggtctac aattgcttgt 2450

caattctgaa gctatgaaga ataaatgttt cagctttgga ttatgaaacc 2500

ccgtttatga tttttaaaac acacttgaaa taaaaatgat tgaactaaat 2550

tttggtccgg tgacatcatt ttgcactgca tagcccatta tacattgtaa 2600

gtctttttgt tttaatttat tactgagagt tttctgtgaa gctacagcat 2650

atctaatcag agaatttctg atttgtttcc ttatactgtg attatgttat 2700

attaaggcat ttgtagtgca gcttaagact gaatttatgc tttttgtaaa 2750

catgatcggt gtaaatgtct tataaacatt ctggagtatg tataacttta 2800

atgaatgaaa tttaatggac ctgattaaaa taaagggatt taatctttgt 2850

taaagtgaag ttagtggaat aaattaccta ctggaatata gtccaagtca 2900

ctaaaggttt aatatttgca tttttgtgct tttattttct ccttccattc 2950

ataattatat acttgaaagt acatctgtag cctatgactt gagtctcttg 3000

aacttctagg aagaggcaaa ctacaaacta ctaggattct gatttcagat 3050

ataggcattc cagaatcttc tctttacgag ttcacctgct agtataatct 3100

ccacaacttg aatggccttg gttgttctgt aattgctgcc aaaatcatca 3150

caagctgtac gtcatcaagg ctccctttgc actcccaaga agaactgttc 3200

attttaaaca aaagtgtatg tctttatttg tattggaaaa tattgtcttt 3250

aaattgtttc ttgttgacac ttcccacaat ggaaaaatta ccaaattaaa 3300

cctgttttat ggatggcagc ttggagcata gcaagaattt ggaggatttg 3350

aattccattc ccagttctca ttgcgttttg tttcttaaaa ctataataat 3400

tggttactgt tataaagttt aaaaggtggt tttaatatga atagcaaatt 3450

ctggtatatt gtgactaacg cttaagaatg cctgtctttg agaggaaggt 3500

gttataatat taatcaacag tgccaaatac actgtgcata tctacaattt 3550

aatctttgaa tgtttgttac tggattagct ccctcctcct tctgtgtgat 3600

ggtaccatgc atagagtcaa tcaaatcctt gtgatgtttt gtatggactt 3650

tgacaacatg taactaatgt gtaaagcaag tttttatgat taaggaatca 3700

aatttattga attttattat tgaaagttga aacttaacat gtataaacaa 3750

aaaacaataa aataataaac tattttcatt gactataaaa aaaagcccac 3800

gcgtccgcc 3809

<210> 20
<211> 1048
<212> DNA
<213> Homo sapiens

<400> 20
gaggtgaggt gttggcagtg gaaaggggtt cgggctcggg gggcgggggg 50

acgcggagcg atggcccgcg ccggccgcag gggcggataa aaagccgtcg 100

cgctgcggga gtgggcggga gggagagggg gtgtccgagg gccacaagag 150

tatgacgggg ctgtacgagc tggtgtggcg ggtgctgcac gcgctgctct 200

gtctgcaccg cacgctcacc tcctggctcc gcgttcggtt cggcacctgg 250

aactggatct ggcggcgctg ctgccgcgcc gcctctgccg cggtcctagc 300

gccgctcggc ttcacgctcc gcaagccccc ggcagtcggc aggaaccgcc 350

gtcaccaccg gcacccgcgc ggggggtcgt gcctggcagc cgcacaccac 400

cggatgcgct ggcgcgcgga cggtcgttcc ttggagaagc tgcctgtgca 450

tatgggcctg gtgatcaccg aggtggagca ggaacccagc ttctcggaca 500

tcgcgagcct cgtggtgtgg tgtatggccg tgggcatctc ctacattagc 550

gtctacgacc accaaggtat tttcaaaaga aataattcca gattgatgga 600

tgaaatttta aaacaacagc aagaacttct gggcctagat tgttcaaaat 650

actcaccaga atttgcaaat agtaatgaca aagatgatca agttttaaat 700

tgccatttgg cagtgaaggt gctgtctccg gaagatggaa aagcagatat 750

tgtaagagct gctcaggact tttgccagtt agtagcccag aagcaaaaga 800

gacccacaga tttggatgta gatacgttag ccagtttact tagttcaaat 850

ggttgtcctg atcctgattt agtattgaag ttcggtcctg tggacagcac 900

attaggcttt cttccctggc acatcagatt gactgagatt gtctctttgc 950

cttcccacct aaacatcagt tatgaggact ttttctctgc ccttcgtcaa 1000

tatgcagcct gtgaacagcg tctgggaaag tagtggtcat tggttgca 1048

<210> 21
<211> 3332
<212> DNA
<213> Homo sapiens

<400> 21
gggggggcgga accagcctgc acgcgctggc tccgggtgac agccgcgcgc 50

ctcggccagg atctgagtga tgagacgtgt ccccactgag gtgccccaca 100

gcagcaggtg ttgagcatgg gctgagaagc tggaccggca ccaaagggct 150

ggcagaaatg ggcgcctggc tgattcctag gcagttggcg gcagcaagga 200

ggagaggccg cagcttctgg agcagagccg agacgaagca gttctggagt 250

gcctgaacgg ccccctgagc cctacccgcc tggcccacta tggtccagag 300

gctgtgggtg agccgcctgc tgcggcaccg gaaagcccag ctcttgctgg 350

tcaacctgct aacctttggc ctggaggtgt gtttggccgc aggcatcacc 400

tatgtgccgc ctctgctgct ggaagtgggg gtagaggaga agttcatgac 450

catggtgctg ggcattggtc cagtgctggg cctggtctgt gtcccgctcc 500

taggctcagc cagtgaccac tggcgtggac gctatggccg ccgccggccc 550

ttcatctggg cactgtcctt gggcatcctg ctgagcctct ttctcatccc 600

aagggccggc tggctagcag ggctgctgtg cccggatccc aggcccctgg 650

agctggcact gctcatcctg ggcgtggggc tgctggactt ctgtggccag 700

gtgtgcttca ctccactgga ggccctgctc tctgacctct tccgggaccc 750

ggaccactgt cgccaggcct actctgtcta tgccttcatg atcagtcttg 800

ggggctgcct gggctacctc ctgcctgcca ttgactggga caccagtgcc 850

ctggcccccc acctgggcac ccaggaggag tgcctctttg gcctgctcac 900

cctcatcttc ctcacctgcg tagcagccac actgctggtg gctgaggagg 950

cagcgctggg ccccaccgag ccagcagaag ggctgtcggc cccctccttg 1000

tcgccccact gctgtccatg ccgggcccgc ttggctttcc ggaacctggg 1050

cgccctgctt ccccggctgc accagctgtg ctgccgcatg ccccgcaccc 1100

tgcgccggct cttcgtggct gagctgtgca gctggatggc actcatgacc 1150

ttcacgctgt tttacacgga tttcgtgggc gaggggctgt accagggcgt 1200

gcccagagct gagccgggca ccgaggcccg gagacactat gatgaaggcg 1250

ttcggatggg cagcctgggg ctgttcctgc agtgcgccat ctccctggtc 1300

ttctctctgg tcatggaccg gctggtgcag cgattcggca ctcgagcagt 1350

ctatttggcc agtgtggcag ctttccctgt ggctgccggt gccacatgcc 1400

tgtcccacag tgtggccgtg gtgacagctt cagccgccct caccgggttc 1450

accttctcag ccctgcagat cctgccctac acactggcct ccctctacca 1500

ccgggagaag caggtgttcc tgcccaaata ccgaggggac actggaggtg 1550

ctagcagtga ggacagcctg atgaccagct tcctgccagg ccctaagcct 1600

ggagctccct tccctaatgg acacgtgggt gctggaggca gtggcctgct 1650

cccacctcca cccgcgctct gcggggcctc tgcctgtgat gtctccgtac 1700

gtgtggtggt gggtgagccc accgaggcca gggtggttcc gggccggggc 1750

atctgcctgg acctcgccat cctggatagt gccttcctgc tgtcccaggt 1800

ggccccatcc ctgtttatgg gctccattgt ccagctcagc cagtctgtca 1850

ctgcctatat ggtgtctgcc gcaggcctgg gtctggtcgc catttacttt 1900

gctacacagg tagtatttga caagagcgac ttggccaaat actcagcgta 1950

gaaaacttcc agcacattgg ggtggagggc ctgcctcact gggtcccagc 2000

tccccgctcc tgttagcccc atggggctgc cgggctggcc gccagtttct 2050

gttgctgcca aagtaatgtg gctctctgct gccaccctgt gctgctgagg 2100

tgcgtagctg cacagctggg ggctggggcg tccctctcct ctctccccag 2150

tctctagggc tgcctgactg gaggccttcc aaggggttt cagtctggac 2200

ttatacaggg aggccagaag ggctccatgc actggaatgt ggggactctg 2250

caggtggatt acccaggctc agggttaaca gctagcctcc tagttgagac 2300

acacctagag aagggttttt gggagctgaa taaactcagt cacctggttt 2350

cccatctcta agccccttaa cctgcagctt cgtttaatgt agctcttgca 2400

tgggagtttc taggatgaaa cactcctcca tgggatttga acatatgaaa 2450

gttatttgta ggggaagagt cctgaggggc aacacacaag aaccaggtcc 2500

cctcagccca cagcactgtc tttttgctga tccacccccc tcttaccttt 2550

tatcaggatg tggcctgttg gtccttctgt tgccatcaca gagacacagg 2600

catttaaata tttaacttat ttatttaaca aagtagaagg gaatccattg 2650

ctagcttttc tgtgttggtg tctaatattt gggtagggtg ggggatcccc 2700

aacaatcagg tcccctgaga tagctggtca ttgggctgat cattgccaga 2750

atcttcttct cctggggtct ggcccccaa aatgcctaac ccaggacctt 2800

ggaaattcta ctcatcccaa atgataattc caaatgctgt tacccaaggt 2850

tagggtgttg aaggaaggta gagggtgggg cttcaggtct caacggcttc 2900

cctaaccacc cctcttctct tggcccagcc tggttccccc cacttccact 2950

cccctctact ctctctagga ctgggctgat gaaggcactg cccaaaattt 3000

cccctacccc caactttccc ctacccccaa ctttccccac cagctccaca 3050

accctgtttg gagctactgc aggaccagaa gcacaaagtg cggtttccca 3100

agcctttgtc catctcagcc cccagagtat atctgtgctt ggggaatctc 3150

acacagaaac tcaggagcac cccctgcctg agctaaggga ggtcttatct 3200

ctcaggggg tttaagtgcc gtttgcaata atgtcgtctt atttatttag 3250

cggggtgaat attttatact gtaagtgagc aatcagagta taatgtttat 3300

ggtgacaaaa ttaaagcctt tcttatatgt tt 3332

<210> 22
<211> 3768
<212> DNA
<213> Homo sapiens

<400> 22
aacaaaagct ggagctccac cgcggtggcg gccgctctag cccctggttc 50

ggcccacctc tgaaggttcc agaatcgata gtgaattcgt ggggaagcgc 100

agccatggct ctgcggaggc tgggggccgc gctgctgctg ctgccgctgc 150

tcgccgccgt ggaagaaacg ctaatggact ccactacagc gactgctgag 200

ctgggctgga tggtgcatcc tccatcaggg tgggaagagg tgagtggcta 250

cgatgagaac atgaacacga tccgcacgta ccaggtgtgc aacgtgtttg 300

agtcaagcca gaacaactgg ctacggacca agtttatccg gcgccgtggc 350

gcccaccgca tccacgtgga gatgaagttt tcggtgcgtg actgcagcag 400

catccccagc gtgcctggct cctgcaagga gaccttcaac ctctattact 450

atgaggctga ctttgactcg gccaccaaga ccttccccaa ctggatggag 500

aatccatggg tgaaggtgga taccattgca gccgacgaga gcttctccca 550

ggtggacctg ggtggccgcg tcatgaaaat caacaccgag gtgcggagct 600

tcggacctgt gtcccgcagc ggcttctacc tggccttcca ggactatggc 650

ggctgcatgt ccctcatcgc cgtgcgtgtc ttctaccgca agtgcccccg 700

catcatccag aatggcgcca tcttccagga aaccctgtcg ggggctgaga 750

gcacatcgct ggtggctgcc cggggcagct gcatcgccaa tgcggaagag 800

gtggatgtac ccatcaagct ctactgtaac ggggacggcg agtggctggt 850

gcccatcggg cgctgcatgt gcaaagcagg cttcgaggcc gttgagaatg 900

gcaccgtctg ccgaggttgt ccatctggga ctttcaaggc caaccaaggg 950

gatgaggcct gtacccactg tcccatcaac agccggacca cttctgaagg 1000

ggccaccaac tgtgtctgcc gcaatggcta ctacagagca gacctggacc 1050

ccctggacat gccctgcaca accatcccct ccgcgcccca ggctgtgatt 1100

tccagtgtca atgagacctc cctcatgctg gagtggaccc ctccccgcga 1150

ctccggaggc cgagaggacc tcgtctacaa catcatctgc aagagctgtg 1200

gctcgggccg gggtgcctgc acccgctgcg gggacaatgt acagtacgca 1250

ccacgccagc taggcctgac cgagccacgc atttacatca gtgacctgct 1300

ggcccacacc cagtacacct tcgagatcca ggctgtgaac ggcgttactg 1350

accagagccc cttctcgcct cagttcgcct ctgtgaacat caccaccaac 1400

caggcagctc catcggcagt gtccatcatg catcaggtga gccgcaccgt 1450

ggacagcatt accctgtcgt ggtcccagcc agaccagccc aatggcgtga 1500

tcctggacta tgagctgcag tactatgaga aggagctcag tgagtacaac 1550

gccacagcca taaaaagccc caccaacacg gtcaccgtgc agggcctcaa 1600

agccggcgcc atctatgtct tccaggtgcg ggcacgcacc gtggcaggct 1650

acgggcgcta cagcggcaag atgtacttcc agaccatgac agaagccgag 1700

taccagacaa gcatccagga gaagttgcca ctcatcatcg gctcctcggc 1750

cgctggcctg gtcttcctca ttgctgtggt tgtcatcgcc atcgtgtgta 1800

acagaagacg ggggtttgag cgtgctgact cggagtacac ggacaagctg 1850

caacactaca ccagtggcca catgacccca ggcatgaaga tctacatcga 1900

tcctttcacc tacgaggacc ccaacgaggc agtgcgggag tttgccaagg 1950

aaattgacat ctcctgtgtc aaaattgagc aggtgatcgg agcaggggag 2000

tttggcgagg tctgcagtgg ccacctgaag ctgccaggca agagagagat 2050

ctttgtggcc atcaagacgc tcaagtcggg ctacacggag aagcagcgcc 2100

gggacttcct gagcgaagcc tccatcatgg gccagttcga ccatcccaac 2150

gtcatccacc tggagggtgt cgtgaccaag agcacacctg tgatgatcat 2200

caccgagttc atggagaatg gctccctgga ctcctttctc cggcaaaacg 2250

atgggcagtt cacagtcatc cagctggtgg gcatgcttcg gggcatcgca 2300

gctggcatga agtacctggc agacatgaac tatgttcacc gtgacctggc 2350

tgcccgcaac atcctcgtca acagcaacct ggtctgcaag gtgtcggact 2400

ttgggctctc acgctttcta gaggacgata cctcagaccc cacctacacc 2450

agtgccctgg gcggaaagat ccccatccgc tggacagccc cggaagccat 2500

ccagtaccgg aagttcacct cggccagtga tgtgtggagc tacggcattg 2550

tcatgtggga ggtgatgtcc tatggggagc ggccctactg ggacatgacc 2600

aaccaggatg taatcaatgc cattgagcag gactatcggc tgccaccgcc 2650

catggactgc ccgagcgccc tgcaccaact catgctggac tgttggcaga 2700

aggaccgcaa ccaccggccc aagttcggcc aaattgtcaa cacgctagac 2750

aagatgatcc gcaatcccaa cagcctcaaa gccatggcgc ccctctcctc 2800

tggcatcaac ctgccgctgc tggaccgcac gatccccgac tacaccagct 2850

ttaacacggt ggacgagtgg ctggaggcca tcaagatggg gcagtacaag 2900

gagagcttcg ccaatgccgg cttcacctcc tttgacgtcg tgtctcagat 2950

gatgatggag gacattctcc gggttgggct cactttggct ggccaccaga 3000

aaaaaatcct gaacagtatc caggtgatgc gggcgcagat gaaccagatt 3050

cagtctgtgg aggtttgaca ttcacctgcc tcggctcacc tcttcctcca 3100

agccccgccc cctctgcccc acgtgccggc cctcctggtg ctctatccac 3150

tgcagggcca gccactcgcc aggaggccac gggccacggg aagaaccaag 3200

cggtgccagc cacgagacgt caccaagaaa acatgcaact caaacgacgg 3250

aaaaaaaaag ggaatgggaa aaaagaaaac agatcctggg agggggcggg 3300

aaatacaagg aatatttttt aaagaggatt ctcataagga aagcaatgac 3350

tgttcttgcg ggggataaaa aagggcttgg gagattcatg cgatgtgtcc 3400

aatcggagac aaaagcagtt tctctccaac tccctctggg aaggtgacct 3450

ggccagagcc aagaaacact ttcagaaaaa caaatgtgaa ggggagagac 3500

aggggccgcc cttggctcct gtccctgctg ctcctctagg cctcactcaa 3550

caaccaagcg cctggaggac gggacagatg gacagacagc caccctgaga 3600

acccctctgg gaaaatctat tcctgccacc actgggcaaa cagaagaatt 3650

tttctgtctt tggagagtat tttagaaact ccaatgaaag acactgtttc 3700

tcctgttggc tcacagggct gaaaggggct tttgtcctcc tgggtcaggg 3750

agaacgcggg gaccccag 3768

<210> 23
<211> 3800
<212> DNA
<213> Homo sapiens

<400> 23
cgggaagcgc agccatggct ctgcggaggc tgggggccgc gctgctgctg 50

ctgccgctgc tcgccgccgt ggaagaaacg ctaatggact ccactacagc 100

gactgctgag ctgggctgga tggtgcatcc tccatcaggg tgggaagagg 150

tgagtggcta cgatgagaac atgaacacga tccgcacgta ccaggtgtgc 200

aacgtgtttg agtcaagcca gaacaactgg ctacggacca agtttatccg 250

gcgccgtggc gcccaccgca tccacgtgga gatgaagttt tcggtgcgtg 300

actgcagcag catccccagc gtgcctggct cctgcaagga gaccttcaac 350

ctctattact atgaggctga ctttgactcg gccaccaaga ccttccccaa 400

ctggatggag aatccatggg tgaaggtgga taccattgca gccgacgaga 450

gcttctccca ggtggacctg ggtggccgcg tcatgaaaat caacaccgag 500

gtgcggagct tcggacctgt gtcccgcagc ggcttctacc tggccttcca 550

ggactatggc ggctgcatgt ccctcatcgc cgtgcgtgtc ttctaccgca 600

agtgcccccg catcatccag aatggcgcca tcttccagga aaccctgtcg 650

ggggctgaga gcacatcgct ggtggctgcc cggggcagct gcatcgccaa 700

tgcggaagag gtggatgtac ccatcaagct ctactgtaac ggggacggcg 750

agtggctggt gcccatcggg cgctgcatgt gcaaagcagg cttcgaggcc 800

gttgagaatg gcaccgtctg ccgaggttgt ccatctggga ctttcaaggc 850

caaccaaggg gatgaggcct gtacccactg tcccatcaac agccggacca 900

cttctgaagg ggccaccaac tgtgtctgcc gcaatggcta ctacagagca 950

gacctggacc ccctggacat gccctgcaca accatcccct ccgcgcccca 1000

ggctgtgatt tccagtgtca atgagacctc cctcatgctg gagtggaccc 1050

ctccccgcga ctccggaggc cgagaggacc tcgtctacaa catcatctgc 1100

aagagctgtg gctcgggccg gggtgcctgc acccgctgcg gggacaatgt 1150

acagtacgca ccacgccagc taggcctgac cgagccacgc atttacatca 1200

gtgacctgct ggcccacacc cagtacacct tcgagatcca ggctgtgaac 1250

ggcgttactg accagagccc cttctcgcct cagttcgcct ctgtgaacat 1300

caccaccaac caggcagctc catcggcagt gtccatcatg catcaggtga 1350

gccgcaccgt ggacagcatt accctgtcgt ggtcccagcc ggaccagccc 1400

aatggcgtga tcctggacta tgagctgcag tactatgaga aggagctcag 1450

tgagtacaac gccacagcca taaaaagccc caccaacacg gtcaccgtgc 1500

agggcctcaa agccggcgcc atctatgtct tccaggtgcg ggcacgcacc 1550

gtggcaggct acgggcgcta cagcggcaag atgtacttcc agaccatgac 1600

agaagccgag taccagacaa gcatccagga gaagttgcca ctcatcatcg 1650

gctcctcggc cgctggcctg gtcttcctca ttgctgtggt tgtcatcgcc 1700

atcgtgtgta acagacgggg gtttgagcgt gctgactcgg agtacacgga 1750

caagctgcaa cactacacca gtggccacat gaccccaggc atgaagatct 1800

acatcgatcc tttcacctac gaggacccca acgaggcagt gcgggagttt 1850

gccaaggaaa ttgacatctc ctgtgtcaaa attgagcagg tgatcggagc 1900

aggggagttt ggcgaggtct gcagtggcca cctgaagctg ccaggcaaga 1950

gagagatctt tgtggccatc aagacgctca agtcgggcta cacggagaag 2000

cagcgccggg acttcctgag cgaagcctcc atcatgggcc agttcgacca 2050

tcccaacgtc atccacctgg agggtgtcgt gaccaagagc acacctgtga 2100

tgatcatcac cgagttcatg gagaatggct ccctggactc ctttctccgg 2150

caaaacgatg ggcagttcac agtcatccag ctggtgggca tgcttcgggg 2200

catcgcagct ggcatgaagt acctggcaga catgaactat gttcaccgtg 2250

acctggctgc ccgcaacatc ctcgtcaaca gcaacctggt ctgcaaggtg 2300

tcggactttg ggctctcacg ctttctagag gacgatacct cagaccccac 2350

ctacaccagt gccctgggcg gaaagatccc catccgctgg acagccccgg 2400

aagccatcca gtaccggaag ttcacctcgg ccagtgatgt gtggagctac 2450

ggcattgtca tgtgggaggt gatgtcctat ggggagcggc cctactggga 2500

catgaccaac caggatgtaa tcaatgccat tgagcaggac tatcggctgc 2550

caccgcccat ggactgcccg agcgccctgc accaactcat gctggactgt 2600

tggcagaagg accgcaacca ccggcccaag ttcggccaaa ttgtcaacac 2650

gctagacaag atgatccgca atcccaacag cctcaaagcc atggcgcccc 2700

tctcctctgg catcaacctg ccgctgctgg accgcacgat ccccgactac 2750

accagcttta acacggtgga cgagtggctg gaggccatca agatggggca 2800

gtacaaggag agcttcgcca atgccggctt cacctccttt gacgtcgtgt 2850

ctcagatgat gatggaggac attctccggg ttggggtcac tttggctggc 2900

caccagaaaa aaatcctgaa cagtatccag gtgatgcggg cgcagatgaa 2950

ccagattcag tctgtggagg tttgacattc acctgcctcg gctcacctct 3000

tcctccaagc cccgcccccc ctgccccacg tgccggccct cctggtgctc 3050

tatccactgc agggccagcc actcgccagg aggccacggg ccacgggaag 3100

aaccaagcgg tgccagccac gagacgtcac caagaaaaca tgcaactcaa 3150

acgacggaaa aaaaaaggga atgggaaaaa agaaaacaga tcctgggagg 3200

gggcgggaaa tacaaggaat atttttaaa gaggattctc ataaggaaag 3250

caatgactgt tcttgcgggg gataaaaaag ggcttgggag attcatgcga 3300

tgtgtccaat cggagacaaa agcagtttct ctccaactcc ctctgggaag 3350

gtgacctggc cagagccaag aaacactttc agaaaaacaa atgtgaaggg 3400

gagagacagg ggccgccctt ggctcctgtc cctgctgctc ctctaggcct 3450

cactcaacaa ccaagcgcct ggaggacggg acagatggac agacagccac 3500

cctgagaacc cctctgggaa aatctattcc tgccaccact gggcaaacag 3550

aagaattttt ctgtctttgg agagtatttt agaaactcca atgaaagaca 3600

ctgtttctcc tgttggctca cagggctgaa aggggctttt gtcctcctgg 3650

gtcagggaga acgcggggac cccagaaagg tcagccttcc tgaggatggg 3700

caacccccag gtctgcagct ccaggtacat atcacgcgca cagcctggca 3750

gcctggccct cctggtgccc actcccgcca gcccctgcct cgaggtcgac 3800

<210> 24
<211> 853
<212> DNA
<213> Homo sapiens

<400> 24
ccgccagctc cttcggcaat gaacttctcc accagcagca gcagcttcgc 50

ctacgaccgg gagttcctcc gcaccctgcc cggcttcctc atcgtggccg 100

agatcgttct ggggctgctg gtatggacgc ttattgctgg aactgagtac 150

ttccgggtcc ccgcatttgg ctgggtcatg tttgtagctg tattttactg 200

ggtcctcacc gtcttcttcc tcattatcta cataacaatg acctacacca 250

ggattcccca ggtgccctgg acaacagtgg gcctgtgctt taacggcagt 300

gccttcgtct tgtacctctc tgccgctgtt gtagatgcat cttccgtctc 350

ccctgagagg gacagtcaca acttcaacag ctgggcggcc tcatcgttct 400

ttgccttcct ggtcaacatc tgctacgctg gaaatacata tttcagtttt 450

atagcatgga gatccaggac catacagtga tttaccattt tgataattaa 500

aaggaaaaaa aaaggaagac tctcactgta aaaacagctg taggtataat 550

gtatattccc agagaattgt atttaactaa ttaatgtttt ttatattctt 600

aaatttgctc acaaattgtg gtttgttaca attaaactgg atacttattt 650

gcaaagtgtt gtagcttata atgaactctt aagtatctta ttaatgtatt 700

aatgtcttca tagatcatat tttcttagac aatgtttaaa tagataaatt 750

gctaatattg agaatgtgtc aagtttgtaa acctaacttt taagatgcca 800

gattcttttt tgattaaatg ttgcaaaatc ccaaaaaaaa aaaaaaaaaa 850

aaa 853

<210> 25
<211> 485
<212> DNA
<213> Homo sapiens

<400> 25
gcccttggca atgaacttct ccaccagcag cagcagcttc gcctacgacc 50

gggagttcct ccgcaccctg cccggcttcc tcatcgtggc cgagatcgtt 100

ctggggctgc tggtatggac gcttattgct ggaactgagt acttccgggt 150

ccccgcattt ggctgggtca tgtttgtagc tgtattttac tgggtcctca 200

ccgtcttctt cctcattatc tacataacaa tgacctacac caggattccc 250

caggtgccct ggacaacagt gggcctgtgc tttaacggca gtgccttcgt 300

cttgtacctc tctgccgctg ttgtagatgc atcttccgtc tcccctgaga 350

gggacagtca caacttcaac agctgggcgg cctcatcgtt ctttgccttc 400

ctggtcacca tctgctacgc tggaaataca tatttcagtt ttatagcatg 450

gagatccagg accatacagt gatttaccaa agggc 485

<210> 26
<211> 2257
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 2159, 2231
<223> unknown base

<400> 26
ggccctgccc tgaaggctgg tcacttgcag aggtaaactc ccctctttga 50

cttctggcca gggtttgtgc tgagctggct gcagccgctc tcagcctcgc 100

tccgggcacg tcgggcagcc tcgggccctc ctgcctgcag gatcatgccc 150

accaccgtgg acgatgtcct ggagcatgga ggggagtttc acttttcca 200

gaagcaaatg tttttcctct tggctctgct ctcggctacc ttcgcgccca 250

tctacgtggg catcgtcttc ctgggcttca cccctgacca ccgctgccgg 300

agccccggag tggccgagct gagtctgcgc tgcggctgga gtcctgcaga 350

ggaactgaac tacacggtgc cgggcccagg acctgcgggc gaagcctccc 400

caagacagtg taggcgctac gaggtggact ggaaccagag cacctttgac 450

tgcgtggacc ccctggccag cctggacacc aacaggagcc gcctgccact 500

gggcccctgc cgggacggct gggtgtacga gacgcctggc tcgtccatcg 550

tcaccgagtt taacctggta tgtgccaact cctggatgtt ggacctattc 600

cagtcatcag tgaatgtagg attctttatt ggctctatga gtatcggcta 650

catagcagac aggtttggcc gtaagctctg cctcctaact acagtcctca 700

taaatgctgc agctggagtt ctcatggcca tttccccaac ctatacgtgg 750

atgttaattt ttcgcttaat ccaaggactg gtcagcaaag caggctggtt 800

aataggctac atcctgatta cagaatttgt tgggcggaga tatcggagaa 850

cagtggggat tttttaccaa gttgcctata cagttgggct cctggtgcta 900

gctggggtgg cttacgcact tcctcactgg aggtggttgc agttcacagt 950

tgctctgccc aacttcttct tcttgctcta ttactggtgc atacctgagt 1000

ctcccaggtg gctgatctcc cagaataaga atgctgaagc catgagaatc 1050

attaagcaca tcgcaaagaa aaatggaaaa tctctacccg cctcccttca 1100

gcgcctgaga cttgaagagg aaactggcaa gaaattgaac ccttcatttc 1150

ttgacttggt cagaactcct cagataagga aacatactat gatattgatg 1200

tacaactggt tcacgagctc tgtgctctac cagggcctca tcatgcacat 1250

gggccttgca ggtgacaata tctacctgga tttcttctac tctgccctgg 1300

ttgaattccc agctgccttc atgatcatcc tcaccatcga ccgcatcgga 1350

cgccgttacc cttgggctgc atcaaatatg gttgcagggg cagcctgtct 1400

ggcctcagtt tttatacctg gtgatctaca atggctaaaa attattatct 1450

catgcttggg aagaatgggg atcacaatgg cctatgagat agtctgcctg 1500

gtcaatgctg agctgtaccc cacattcatt aggaatcttg gcgtccacat 1550

ctgttcctca atgtgtgaca ttggtggcat catcacgcca ttcctggtct 1600

accggctcac taacatctgg cttgagctcc cgctgatggt tttcggcgta 1650

cttggcttgg ttgctggagg tctggtgctg ttgcttccag aaactaaagg 1700

gaaagctttg cctgagacca tcgaggaagc cgaaaatatg caaagaccaa 1750

gaaaaaataa agaaaagatg atttacctcc aagttcagaa actagacatt 1800

ccattgaact aagaagagag accgttgctg ctgtcatgac ctagctttga 1850

tggcagcaag accaaaagta gaaatccctg cactcatcac aaagcccata 1900

caactcaacc aaacttaccc ctgagcccta tcaacctagg tctacagcca 1950

gtggagtcta ttgtacactg tggaaaaata cccatgggac cagatcctgc 2000

caaattcttc cagctcactt tattctcagc attcctagga cattggacat 2050

tggttttctg gagggttttt tttccgatct ttgtattttt ttaaatttga 2100

ttcttttctt tgcaatgcta gcaaccagaa tacatagggg aactgtgggc 2150

taggcaaana aaatagaaaa agtgtgaaaa acagtaaagt tgggagagga 2200

gcatctattt tcttaaagaa ataaaacacc naaaacaaaa aaaaaaaaaa 2250

aaaaaaa 2257

<210> 27
<211> 1725
<212> DNA
<213> Homo sapiens

<400> 27
gcccttctcc tgcctgcagg atcatgccca ccaccgtgga cgatgtcctg 50

gagcatggag gggagtttca ctttttccag aagcaaatgt ttttcctctt 100

ggctctgctc tcggctacct tcgcgcccat ctacgtgggc atcgtcttcc 150

tgggcttcac ccctgaccac cgctgccgga gccccggagt ggccgagctg 200

agtctgcgct gcggctggag tcctgcagag gaactgaact acacggtgcc 250

gggcccagga cctgcgggcg aagcctcccc aagacagtgt aggcgctacg 300

aggtggactg gaaccagagc accttcgact gcgtggaccc cctggccagc 350

ctggacacca acaggagccg cctgccactg ggcccctgcc gggacggctg 400

ggtgtacgag acgcctggct cgtccatcgt caccgagttt aacctggtat 450

gtgccaactc ctggatgttg gacctattcc agtcatcagt gaatgtagga 500

ttctttattg gctctatgag tatcggctac atagcagaca ggtttggccg 550

taagctctgc ctcctaacta cagtcctcat aaatgctgca gctggagttc 600

tcatggccat ttccccaacc tatacgtgga tgttaatttt tcgcttaatc 650

caaggactgg tcagcaaagc aggctggtta ataggctaca tcctgattac 700

agaatttgtt gggcggagat atcggagaac agtggggatt ttttaccaag 750

ttgcctatac agttgggctc ctggtgctag ctggggtggc ttacgcactt 800

cctcactgga ggtggttgca gttcacagtt tctctgccca acttcttctt 850

cttgctctat tactggtgca tacctgagtc tcccaggtgg ctgatctccc 900

agaataagaa tgctgaagcc atgagaatca ttaagcacat cgcaaagaaa 950

aatggaaaat ctctacccgc ctcccttcag cgcctgagac ttgaagagga 1000

aactggcaag aaattgaacc cttcatttct tgacttggtc agaactcctc 1050

agataaggaa acatactatg atattgatgt acaactggtt cacgagctct 1100

gtgctctacc agggcctcat catgcacatg ggccttgcag gtgacaatat 1150

ctacctggat ttcttctact ctgccctggt tgaattccca gctgccttca 1200

tgatcatcct caccatcgac cgcatcggac gccgttaccc ttgggctgca 1250

tcaaatatgg ttgcaggggc agcctgtctg gcctcagttt ttatacctgg 1300

tgatctacaa tggctaaaaa ttattatctc atgcttggga agaatgggga 1350

tcacaatggc ctatgagata gtctgcctgg tcaatgctga gctgtacccc 1400

acattcatta ggaatcttgg cgtccacatc tgttcctcaa tgtgtgacat 1450

tggtggcatc atcacgccat tcctggtcta ccggctcact aacatctggc 1500

ttgagctccc gctgatggtt ttcggcgtgc ttggcttggt tgctggaggt 1550

ctggtgctgt tgcttccaga aactaaaggg aaagctttgc ctgagaccat 1600

cgaggaagcc gaaaatatgc aaagaccaag aaaaaataaa gaaaagatga 1650

tttacctcca agttcagaaa ctagacattc cattgaacta agaagagaga 1700

ccgttgctgc tgtcatgaca agggc 1725

<210> 28
<211> 2280
<212> DNA
<213> Homo sapiens

<400> 28
cgggccaggt ttccaggctc ggccgccgcc tccatcccag cacctgcgga 50

gggagcgctg accatggctc cctggcctga attgggagat gcccagccca 100

accccgataa gtacctcgaa ggggccgcag gtcagcagcc cactgcccct 150

gataaaagca aagagaccaa caaaacagat aacactgagg cacctgtaac 200

caagattgaa cttctgccgt cctactccac ggctacactg atagatgagc 250

ccactgaggt ggatgacccc tggaacctac ccactcttca ggactcgggg 300

atcaagtggt cagagagaga caccaaaggg aagattctct gtttcttcca 350

agggattggg agattgattt tacttctcgg atttctctac tttttcgtgt 400

gctccctgga tattcttagt agcgccttcc agctggttgg aggaaaaatg 450

gcaggacagt tcttcagcaa cagctctatt atgtccaacc ctttgttggg 500

gctggtgatc ggggtgctgg tgaccgtctt ggtgcagagc tccagcacct 550

caacgtccat cgttgtcagc atggtgtcct cttcattgct cactgttcgg 600

gctgccatcc ccattatcat gggggccaac attggaacgt caatcaccaa 650

cactattgtt gcgctcatgc aggtgggaga tcggagtgag ttcagaagag 700

cttttgcagg agccactgtc catgacttct tcaactggct gtccgtgttg 750

gtgctcttgc ccgtggaggt ggccacccat tacctcgaga tcataaccca 800

gcttatagtg gagagcttcc acttcaagaa tggagaagat gccccagatc 850

ttctgaaagt catcactaag cccttcacaa agctcattgt ccagctggat 900

aaaaaagtta tcagccaaat tgcaatgaac gatgaaaaag cgaaaaacaa 950

gagtcttgtc aagatttggt gcaaaacttt taccaacaag acccagatta 1000

acgtcactgt tccctcgact gctaactgca cctccccttc cctctgttgg 1050

acggatggca tccaaaactg gaccatgaag aatgtgacct acaaggagaa 1100

catcgccaaa tgccagcata tctttgtgaa tttccacctc ccggatcttg 1150

ctgtgggcac catcttgctc atactctccc tgctggtcct ctgtggttgc 1200

ctgatcatga ttgtcaagat cctgggctct gtgctcaagg ggcaggtcgc 1250

cactgtcatc aagaagacca tcaacactga tttccccttt ccctttgcat 1300

ggttgactgg ctacctggcc atcctcgtcg gggcaggcat gaccttcatc 1350

gtacagagca gctctgtgtt cacgtcggcc ttgacccccc tgattggaat 1400

cggcgtgata accattgaga gggcttatcc actcacgctg ggctccaaca 1450

tcggcaccac caccaccgcc atcctggccg ccttagccag ccctggcaat 1500

gcattgagga gttcactcca gatcgccctg tgccactttt tcttcaacat 1550

ctccggcatc ttgctgtggt acccgatccc gttcactcgc ctgcccatcc 1600

gcatggccaa ggggctgggc aacatctctg ccaagtatcg ctggttcgcc 1650

gtcttctacc tgatcatctt cttcttcctg atcccgctga cggtgtttgg 1700

cctctcgctg gccggctggc gggtgctggt tggtgtcggg gttcccgtcg 1750

tcttcatcat catcctggta ctgtgcctcc gactcctgca gtctcgctgc 1800

ccacgcgtcc tgccgaagaa actccagaac tggaacttcc tgccgctgtg 1850

gatgcgctcg ctgaagccct gggatgccgt cgtctccaag ttcaccggct 1900

gcttccagat gcgctgctgc tactgctgcc gcgtgtgctg ccgcgcgtgc 1950

tgcttgctgt gtggctgccc caagtgctgc cgctgcagca agtgctgcga 2000

ggacttggag gaggcgcagg aggggcagga tgtccctgtc aaggctcctg 2050

agacctttga taacataacc attagcagag aggctcaggg tgaggtccct 2100

gcctcggact caaagaccga atgcacggcc ttgtagggga cgccccagat 2150

tgtcagggat ggggggatgg tccttgagtt ttgcatgctc tcctccctcc 2200

cacttctgca ccctttcacc acctcgagga gatttgctcc ccattagcga 2250

atgaaattga tgcagtccta aaaaaaaaaa 2280

<210> 29
<211> 2366
<212> DNA
<213> Homo sapiens

<400> 29
cagcccagca cctgcggagg gagcgctgac catggctccc tggcctgaat 50

tgggagatgc ccagcccaac cccgataagt acctcgaagg ggccgcaggt 100

cagcagccca ctgcccctga taaaagcaaa gagaccaaca aaacagataa 150

cactgaggca cctgtaacca agattgaact tctgccgtcc tactccacgg 200

ctacactgat agatgagccc actgaggtgg atgacccctg gaacctaccc 250

actcttcagg actcggggat caagtggtca gagagagaca ccaaagggaa 300

gattctctgt ttcttccaag ggattgggag attgatttta cttctcggat 350

ttctctactt tttcgtgtgc tccctggata ttcttagtag cgccttccag 400

ctggttggag gaaaaatggc aggacagttc ttcagcaaca gctctattat 450

gtccaaccct ttgttggggc tggtgatcgg ggtgctggtg accgtcttgg 500

tgcagagctc cagcacctca acgtccatcg ttgtcagcat ggtgtcctct 550

tcattgctca ctgttcgggc tgccatcccc attatcatgg gggccaacat 600

tggaacgtca atcaccaaca ctattgttgc gctcatgcag gtgggagatc 650

ggagtgagtt cagaagagct tttgcaggag ccactgtcca tgacttcttc 700

aactggctgt ccgtgttggt gctcttgccc gtggaggtgg ccacccatta 750

cctcgagatc ataacccagc ttatagtgga gagcttccac ttcaagaatg 800

gagaagatgc cccagatctt ctgaaagtca tcactaagcc cttcacaaag 850

ctcattgtcc agctggataa aaaagttatc agccaaattg caatgaacga 900

tgaaaaagcg aaaaacaaga gtcttgtcaa gatttggtgc aaaactttta 950

ccaacaagac ccagattaac gtcactgttc cctcgactgc taactgcacc 1000

tccccttccc tctgttggac ggatggcatc caaaactgga ccatgaagaa 1050

tgtgacctac aaggagaaca tcgccaaatg ccagcatatc tttgtgaatt 1100

tccacctccc ggatcttgct gtgggcacca tcttgctcat actctccctg 1150

ctggtcctct gtggttgcct gatcatgatt gtcaagatcc tgggctctgt 1200

gctcaagggg caggtcgcca ctgtcatcaa gaagaccatc aacactgatt 1250

tccccttcc ctttgcatgg ttgactggct acctggccat cctcgtcggg 1300

gcaggcatga ccttcatcgt acagagcagc tctgtgttca cgtcggcctt 1350

gacccccctg attggaatcg gcgtgataac cattgagagg gcttatccac 1400

tcacgctggg ctccaacatc ggcaccacca ccaccgccat cctggccgcc 1450

ttagccagcc ctggcaatgc attgaggagt tcactccaga tcgccctgtg 1500

ccacttttc ttcaacatct ccggcatctt gctgtggtac ccgatcccgt 1550

tcactcgcct gcccatccgc atggccaagg ggctgggcaa catctctgcc 1600

aagtatcgct ggttcgccgt cttctacctg atcatcttct tcttcctgat 1650

cccgctgacg gtgtttggcc tctcgctggc cggctggcgg gtgctggttg 1700

gtgtcggggt tcccgtcgtc ttcatcatca tcctggtact gtgcctccga 1750

ctcctgcagt ctcgctgccc acgcgtcctg ccgaagaaac tccagaactg 1800

gaacttcctg ccgctgtgga tgcgctcgct gaagccctgg gatgccgtcg 1850

tctccaagtt caccggctgc ttccagatgc gctgctgctg ctgctgccgc 1900

gtgtgctgcc gcgcgtgctg cttgctgtgt ggctgcccca agtgctgccg 1950

ctgcagcaag tgctgcgagg acttggagga ggcgcaggag gggcaggatg 2000

tccctgtcaa ggctcctgag acctttgata acataaccat tagcagagag 2050

gctcagggtg aggtccctgc ctcggactca aagaccgaat gcacggcctt 2100

gtaggggacg ccccagattg tcagggatgg ggggatggtc cttgagtttt 2150

gcatgctctc ctccctccca cttctgcacc ctttcaccac ctcgaggaga 2200

tttgctcccc attagcgaat gaaattgatg cagtcctacc taactcgatt 2250

ccctttggct tggtgggtag gcctgcaggg cacttttatt ccaacccatg 2300

gcctccatga ctttttcaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2350

aaaaaaaaaa aaaaaa 2366

<210> 30
<211> 1633
<212> DNA
<213> Homo sapiens

<400> 30
ggcgtgggag ctgcctaggc cgggccctgc cagggagcaa gtctgcatcg 50

cagaggccag ggcagagtgg ctcctcacag cctgaagctc atccttctgc 100

acgggccagc caggccagca cagaggcacc agggcagcag tgcacacagg 150

tccccgggga ccccaccatg tggagcggat ggtggctgtg gcccccttgtg 200

gccgtctgca ctgcagactt ctttcgggac gaggcagaga ggatcatgag 250

ggactcccct gtcattgatg ggcacaatga cctcccctgg cagctgctgg 300

atatgttcaa caaccggctg caggacgaga gggccaacct gaccaccttg 350

gccggcacac acaccaacat ccccaagctg agggccggct ttgtgggagg 400

ccagttctgg tccgtgtaca cgccctgcga cacccagaac aaagacgccg 450

tgcggaggac gctggagcag atggacgtgg tccaccgcat gtgccggatg 500

tacccggaga ccttcctgta tgtcaccagc agtgcaggca ttcggcaggc 550

cttccgggaa gggaaggtgg ccagcctgat cggcgtggag ggcggccact 600

ccattgacag cagtttgggc gtcctgcggg cactctatca gctgggcatg 650

cggtacctga ccctcaccca cagctgcaac acgccctggg ctgacaactg 700

gctggtggac acgggagaca gcgagcccca gagccaaggc ttgtcaccct 750

ttgggcagcg tgtggtgaag gagctgaacc gtctgggggt cctcatcgac 800

ttggctcacg tgtctgtggc caccatgaag gccaccctgc agctgtccag 850

agccccggtc atcttcagcc actcctcggc ctacagcgtg tgcgcaagcc 900

ggcgcaacgt gcctgacgac gtcctgaggc tggtgaaaca gacagacagc 950

ctggtgatgg tgaacttcta caacaattac atttcctgca ccaacaaggc 1000

caacctgtcc caagtggccg accatctgga tcacatcaag gaggtggcag 1050

gagccagagc cgtgggtttt ggtggggact ttgatggtgt tccaagggtc 1100

cctgaggggc tggaggacgt ctccaagtat ccagacctga tcgctgagct 1150

gctcaggagg aactggacgg aggcggaggt caagggcgca ctggctgaca 1200

acctgctgag ggtcttccag gctgtggaac aggccagcaa cctcacacag 1250

gctcccgagg aggagcccat cccgctggac cagctgggtg gctcctgcag 1300

gacccattac ggctactcct ctggggcttc cagcctccat cgccactggg 1350

ggctcctgct ggcctccctc gctcccctgg tcctctgtct gtctctcctg 1400

tgaaacctgg gagaccagag tccccttttag ggttcccgga gctccgggaa 1450

gacccgccca tcccaggact ccagatgcca ggagccctgc tgcccacatg 1500

caaggaccag catctcctga gaggacgcct gggcttacct ggggggcagg 1550

atgcctgggg acagttcagg acacacacac agtaggcccg caataaaagc 1600

aacacccctt caaaaaaaaa aaaaaaaaaa aaa 1633

<210> 31
<211> 1589
<212> DNA
<213> Homo sapiens

<400> 31
catctaaagc ctcctcagcc ttctgagtca gcctgaaagg aacaggccga 50

actgctgtat gggctctact gccagtgtga cctcaccctc tccagtcacc 100

cctcctcagt tccagctatg agttcctgca acttcacaca tgccaccttt 150

gtgcttattg gtatcccagg attagagaaa gcccatttct gggttggctt 200

ccccctcctt tccatgtatg tagtggcaat gtttggaaac tgcatcgtgg 250

tcttcatcgt aaggacggaa cgcagcctgc acgctccgat gtacctcttt 300

ctctgcatgc ttgcagccat tgacctggcc ttatccacat ccaccatgcc 350

taagatcctt gccctttttct ggtttgattc ccgagagatt agctttgagg 400

cctgtcttac ccagatgttc tttattcatg ccctctcagc cattgaatcc 450

accatcctgc tggccatggc ctttgaccgt tatgtggcca tctgccaccc 500

actgcgccat gctgcagtgc tcaacaatac agtaacagcc cagattggca 550

tcgtggctgt ggtccgcgga tccctctttt ttttcccact gcctctgctg 600

atcaagcggc tggccttctg ccactccaat gtcctctcgc actcctattg 650

tgtccaccag gatgtaatga agttggccta tgcagacact ttgcccaatg 700

tggtatatgg tcttactgcc attctgctgg tcatgggcgt ggacgtaatg 750

ttcatctcct tgtcctattt tctgataata cgaacggttc tgcaactgcc 800

ttccaagtca gagcgggcca aggcctttgg aacctgtgtg tcacacattg 850

gtgtggtact cgccttctat gtgccactta ttggcctctc agttgtacac 900

cgctttggaa acagccttca tcccattgtg cgtgttgtca tgggtgacat 950

ctacctgctg ctgcctcctg tcatcaatcc catcatctat ggtgccaaaa 1000

ccaaacagat cagaacacgg gtgctggcta tgttcaagat cagctgtgac 1050

aaggacttgc aggctgtggg aggcaagtga cccttaacac tacacttctc 1100

cttatcttta ttggcttgat aaacataatt atttctaaca ctagcttatt 1150

tccagttgcc cataagcaca tcagtacttt tctctggctg gaatagtaaa 1200

ctaaagtatg gtacatctac ctaaaggact attatgtgga ataatacata 1250

ctaatgaagt attacatgat ttaaagacta caataaaacc aaacatgctt 1300

ataacattaa gaaaaacaat aaagatacat gattgaaacc aagttgaaaa 1350

atagcatatg ccttggagga aatgtgctca aattactaat gatttagtgt 1400

tgtccctact ttctctctct ttttttcttc ttttttttt attatggtta 1450

gctgtcacat acaacttttt tttttttga gatggggtct ccagcctggg 1500

caacagagca agaccctgtc tcaaagcata aaatggaata acatatcaaa 1550

tgaaacaggg aaaatgaagc tgacaattta tgggagcca 1589

<210> 32
<211> 7941
<212> DNA
<213> Homo sapiens

<400> 32
cacacatacg cacgcacgat ctcacttcga tctatacact ggaggattaa 50

aacaaacaaa caaaaaaaac atttccttcg ctccccctcc ctctccactc 100

tgagaagcag aggagccgca cggcgagggg ccgcagaccg tctggaaatg 150

cgaatcctaa agcgtttcct cgcttgcatt cagctcctct gtgtttgccg 200

cctggattgg gctaatggat actacagaca acagagaaaa cttgttgaag 250

agattggctg gtcctataca ggagcactga atcaaaaaaa ttggggaaag 300

aaatatccaa catgtaatag cccaaaacaa tctcctatca atattgatga 350

agatcttaca caagtaaatg tgaatcttaa gaaacttaaa tttcagggtt 400

gggataaaac atcattggaa aacacattca ttcataacac tgggaaaaca 450

gtggaaatta atctcactaa tgactaccgt gtcagcggag gagtttcaga 500

aatggtgttt aaagcaagca agataacttt tcactgggga aaatgcaata 550

tgtcatctga tggatcagag catagtttag aaggacaaaa atttccactt 600

gagatgcaaa tctactgctt tgatgcggac cgattttcaa gttttgagga 650

agcagtcaaa ggaaaaggga agttaagagc tttatccatt ttgtttgagg 700

ttgggacaga agaaaatttg gatttcaaag cgattattga tggagtcgaa 750

agtgttagtc gttttgggaa gcaggctgct ttagatccat tcatactgtt 800

gaaccttctg ccaaactcaa ctgacaagta ttacatttac aatggctcat 850

tgacatctcc tccctgcaca gacacagttg actggattgt ttttaaagat 900

acagttagca tctctgaaag ccagttggct gttttttgtg aagttcttac 950

aatgcaacaa tctggttatg tcatgctgat ggactactta caaaacaatt 1000

ttcgagagca acagtacaag ttctctagac aggtgttttc ctcatacact 1050

ggaaaggaag agattcatga agcagtttgt agttcagaac cagaaaatgt 1100

tcaggctgac ccagagaatt ataccagcct tcttgttaca tgggaaagac 1150

ctcgagtcgt ttatgatacc atgattgaga agtttgcagt tttgtaccag 1200

cagttggatg gagaggacca aaccaagcat gaatttttga cagatggcta 1250

tcaagacttg ggtgctattc tcaataattt gctacccaat atgagttatg 1300

ttcttcagat agtagccata tgcactaatg gcttatatgg aaaatacagc 1350

gaccaactga ttgtcgacat gcctactgat aatcctgaac ttgatctttt 1400

ccctgaatta attggaactg aagaaataat caaggaggag gaagagggaa 1450

aagacattga agaaggcgct attgtgaatc ctggtagaga cagtgctaca 1500

aaccaaatca ggaaaaagga accccagatt tctaccacaa cacactacaa 1550

tcgcataggg acgaaataca atgaagccaa gactaaccga tccccaacaa 1600

gaggaagtga attctctgga aagggtgatg ttcccaatac atctttaaat 1650

tccacttccc aaccagtcac taaattagcc acagaaaaag atatttcctt 1700

gacttctcag actgtgactg aactgccacc tcacactgtg gaaggtactt 1750

cagcctcttt aaatgatggc tctaaaactg ttcttagatc tccacatatg 1800

aacttgtcgg ggactgcaga atccttaaat acagtttcta taacagaata 1850

tgaggaggag agtttattga ccagtttcaa gcttgatact ggagctgaag 1900

attcttcagg ctccagtccc gcaacttctg ctatcccatt catctctgag 1950

aacatatccc aagggtatat attttcctcc gaaaacccag agacaataac 2000

atatgatgtc cttataccag aatctgctag aaatgcttcc gaagattcaa 2050

cttcatcagg ttcagaagaa tcactaaagg atccttctat ggagggaaat 2100

gtgtggtttc ctagctctac agacataaca gcacagcccg atgttggatc 2150

aggcagagag agctttctcc agactaatta cactgagata cgtgttgatg 2200

aatctgagaa gacaaccaag tccttttctg caggcccagt gatgtcacag 2250

ggtccctcag ttacagatct ggaaatgcca cattattcta cctttgccta 2300

cttcccaact gaggtaacac ctcatgcttt taccccatcc tccagacaac 2350

aggatttggt ctccacggtc aacgtggtat actcgcagac aacccaaccg 2400

gtatacaatg gtgagacacc tcttcaacct tcctacagta gtgaagtctt 2450

tcctctagtc acccctttgt tgcttgacaa tcagatcctc aacactaccc 2500

ctgctgcttc aagtagtgat tcggccttgc atgctacgcc tgtatttccc 2550

agtgtcgatg tgtcatttga atccatcctg tcttcctatg atggtgcacc 2600

tttgcttcca ttttcctctg cttccttcag tagtgaattg tttcgccatc 2650

tgcatacagt ttctcaaatc cttccacaag ttacttcagc taccgagagt 2700

gataaggtgc ccttgcatgc ttctctgcca gtggctgggg gtgatttgct 2750

attagagccc agccttgctc agtattctga tgtgctgtcc actactcatg 2800

ctgcttcaga gacgctggaa tttggtagtg aatctggtgt tctttataaa 2850

acgcttatgt tttctcaagt tgaaccaccc agcagtgatg ccatgatgca 2900

tgcacgttct tcagggcctg aaccttctta tgccttgtct gataatgagg 2950

gctcccaaca catcttcact gtttcttaca gttctgcaat acctgtgcat 3000

gattctgtgg gtgtaactta tcagggttcc ttatttagcg gccctagcca 3050

tataccaata cctaagtctt cgttaataac cccaactgca tcattactgc 3100

agcctactca tgccctctct ggtgatgggg aatggtctgg agcctcttct 3150

gatagtgaat ttcttttacc tgacacagat gggctgacag cccttaacat 3200

ttcttcacct gtttctgtag ctgaatttac atatacaaca tctgtgtttg 3250

gtgatgataa taaggcgctt tctaaaagtg aaataatata tggaaatgag 3300

actgaactgc aaattccttc tttcaatgag atggtttacc cttctgaaag 3350

cacagtcatg cccaacatgt atgataatgt aaataagttg aatgcgtctt 3400

tacaagaaac ctctgtttcc atttctagca ccaagggcat gtttccaggg 3450

tcccttgctc ataccaccac taaggttttt gatcatgaga ttagtcaagt 3500

tccagaaaat aacttttcag ttcaacctac acatactgtc tctcaagcat 3550

ctggtgacac ttcgcttaaa cctgtgctta gtgcaaactc agagccagca 3600

tcctctgacc ctgcttctag tgaaatgtta tctccttcaa ctcagctctt 3650

attttatgag acctcagctt cttttagtac tgaagtattg ctacaacctt 3700

cctttcaggc ttctgatgtt gacaccttgc ttaaaactgt tcttccagct 3750

gtgcccagtg atccaatatt ggttgaaacc cccaaagttg ataaaattag 3800

ttctacaatg ttgcatctca ttgtatcaaa ttctgcttca agtgaaaaca 3850

tgctgcactc tacatctgta ccagtttttg atgtgtcgcc tacttctcat 3900

atgcactctg cttcacttca aggtttgacc atttcctatg caagtgagaa 3950

atatgaacca gttttgttaa aaagtgaaag ttcccaccaa gtggtacctt 4000

ctttgtacag taatgatgag ttgttccaaa cggccaattt ggagattaac 4050

caggcccatc ccccaaaagg aaggcatgta tttgctacac ctgttttatc 4100

aattgatgaa ccattaaata cactaataaa taagcttata cattccgatg 4150

aaatttaac ctccaccaaa agttctgtta ctggtaaggt atttgctggt 4200

attccaacag ttgcttctga tacatttgta tctactgatc attctgttcc 4250

tataggaaat gggcatgttg ccattacagc tgtttctccc cacagagatg 4300

gttctgtaac ctcaacaaag ttgctgtttc cttctaaggc aacttctgag 4350

ctgagtcata gtgccaaatc tgatgccggt ttagtgggtg gtggtgaaga 4400

tggtgacact gatgatgatg gtgatgatga tgatgacaga gatagtgatg 4450

gcttatccat tcataagtgt atgtcatgct catcctatag agaatcacag 4500

gaaaaggtaa tgaatgattc agacacccac gaaaacagtc ttatggatca 4550

gaataatcca atctcatact cactatctga gaattctgaa gaagataata 4600

gagtcacaag tgtatcctca gacagtcaaa ctggtatgga cagaagtcct 4650

ggtaaatcac catcagcaaa tgggctatcc caaaagcaca atgatggaaa 4700

agaggaaaat gacattcaga ctggtagtgc tctgcttcct ctcagccctg 4750

aatctaaagc atgggcagtt ctgacaagtg atgaagaaag tggatcaggg 4800

caaggtacct cagatagcct taatgagaat gagacttcca cagatttcag 4850

ttttgcagac actaatgaaa aagatgctga tgggatcctg gcagcaggtg 4900

actcagaaat aactcctgga ttcccacagt ccccaacatc atctgttact 4950

agcgagaact cagaagtgtt ccacgtttca gaggcagagg ccagtaatag 5000

tagccatgag tctcgtattg gtctagctga ggggttggaa tccgagaaga 5050

aggcagttat acccccttgtg atcgtgtcag ccctgacttt tatctgtcta 5100

gtggttcttg tgggtattct catctactgg aggaaatgct tccagactgc 5150

acacttttac ttagaggaca gtacatcccc tagagttata tccacacctc 5200

caacacctat ctttccaatt tcagatgatg tcggagcaat tccaataaag 5250

cactttccaa agcatgttgc agatttacat gcaagtagtg ggtttactga 5300

agaatttgag acactgaaag agttttacca ggaagtgcag agctgtactg 5350

ttgacttagg tattacagca gacagctcca accacccaga caacaagcac 5400

aagaatcgat acataaatat cgttgcctat gatcatagca gggttaagct 5450

agcacagctt gctgaaaagg atggcaaact gactgattat atcaatgcca 5500

attatgttga tggctacaac agaccaaaag cttatattgc tgcccaaggc 5550

ccactgaaat ccacagctga agatttctgg agaatgatat gggaacataa 5600

tgtggaagtt attgtcatga taacaaacct cgtggagaaa ggaaggagaa 5650

aatgtgatca gtactggcct gccgatggga gtgaggagta cgggaacttt 5700

ctggtcactc agaagagtgt gcaagtgctt gcctattata ctgtgaggaa 5750

ttttactcta agaaacacaa aaataaaaaa gggctcccag aaaggaagac 5800

ccagtggacg tgtggtcaca cagtatcact acacgcagtg gcctgacatg 5850

ggagtaccag agtactccct gccagtgctg acctttgtga gaaaggcagc 5900

ctatgccaag cgccatgcag tggggcctgt tgtcgtccac tgcagtgctg 5950

gagttggaag aacaggcaca tatattgtgc tagacagtat gttgcagcag 6000

attcaacacg aaggaactgt caacatattt ggcttcttaa aacacatccg 6050

ttcacaaaga aattatttgg tacaaactga ggagcaatat gtcttcattc 6100

atgatacact ggttgaggcc atacttagta aagaaactga ggtgctggac 6150

agtcatattc atgcctatgt taatgcactc ctcattcctg gaccagcagg 6200

caaaacaaag ctagagaaac aattccagct cctgagccag tcaaatatac 6250

agcagagtga ctattctgca gccctaaagc aatgcaacag ggaaaagaat 6300

cgaacttctt ctatcatccc tgtggaaaga tcaagggttg gcatttcatc 6350

cctgagtgga gaaggcacag actacatcaa tgcctcctat atcatgggct 6400

attaccagag caatgaattc atcattaccc agcaccctct ccttcatacc 6450

atcaaggatt tctggaggat gatatgggac cataatgccc aactggtggt 6500

tatgattcct gatggccaaa acatggcaga agatgaattt gtttactggc 6550

caaataaaga tgagcctata aattgtgaga gctttaaggt cactcttatg 6600

gctgaagaac acaaatgtct atctaatgag gaaaaactta taattcagga 6650

ctttatctta gaagctacac aggatgatta tgtacttgaa gtgaggcact 6700

ttcagtgtcc taaatggcca aatccagata gccccattag taaaactttt 6750

gaacttataa gtgttataaa agaagaagct gccaataggg atgggcctat 6800

gattgttcat gatgagcatg gaggagtgac ggcaggaact ttctgtgctc 6850

tgacaaccct tatgcaccaa ctagaaaaag aaaattccgt ggatgtttac 6900

caggtagcca agatgatcaa tctgatgagg ccaggagtct ttgctgacat 6950

tgagcagtat cagtttctct acaaagtgat cctcagcctt gtgagcacaa 7000

ggcaggaaga gaatccatcc acctctctgg acagtaatgg tgcagcattg 7050

cctgatggaa atatagctga gagcttagag tctttagttt aacacagaaa 7100

ggggtggggg gactcacatc tgagcattgt tttcctcttc ctaaaattag 7150

gcaggaaaat cagtctagtt ctgttatctg ttgatttccc atcacctgac 7200

agtaactttc atgacatagg attctgccgc caaatttata tcattaacaa 7250

tgtgtgcctt tttgcaagac ttgtaattta cttattatgt ttgaactaaa 7300

atgattgaat tttacagtat ttctaagaat ggaattgtgg tatttttttc 7350

tgtattgatt ttaacagaaa atttcaattt atagaggtta ggaattccaa 7400

actacagaaa atgtttgttt ttagtgtcaa atttttagct gtatttgtag 7450

caattatcag gtttgctaga aatataactt ttaatacagt agcctgtaaa 7500

taaaacactc ttccatatga tattcaacat tttacaactg cagtattcac 7550

ctaaagtaga aataatctgt tacttattgt aaatactgcc ctagtgtctc 7600

catggaccaa atttatattt ataattgtag atttttatat tttactactg 7650

agtcaagttt tctagttctg tgtaattgtt tagtttaatg acgtagttca 7700

ttagctggtc ttactctacc agttttctga cattgtattg tgttacctaa 7750

gtcattaact ttgtttcagc atgtaatttt aacttttgtg gaaaatagaa 7800

ataccttcat tttgaaagaa gtttttatga gaataacacc ttaccaaaca 7850

ttgttcaaat ggtttttatc caaggaattg caaaaataaa tataaatatt 7900

gccattaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa a 7941

<210> 33
<211> 2755
<212> DNA
<213> Homo sapiens

<400> 33
attgtctggg aattgcagcc gcggggcggg cggcggcggc ggcggcggcg 50

gccgggaccc agcgggccag gtggggacgg cgcggagcgg gtgcgggaga 100

tgccgtgcgg gactggggcc acctgagccg cccgcctcgt ccccgccttc 150

tgtgggaagg atgtgcgcgc ggatggccgg tcgcacaaga gcggcccctc 200

gggggcccta cggcccctgg ctctgcctcc tggtggccct cgccctggac 250

gtcgtgagag tggactgtgg ccaggctccc ctggaccctg tctacctgca 300

tgtgacagcc gcccgcccag cccagcccac actctggact gccaagctag 350

accgcttcaa gggctccagg caccacacca ccctcatcac ctgccaccgt 400

gctgggctca cagagccaga ttccagcagt ccccttgaac tgtctgagtt 450

cctatgggtg gactttgtgg tggagaatag cactggtggg ggcgtagcgg 500

tcactcgccc cgtcacgtgg cagctggagt acccaggcca ggcccctgaa 550

gcagagaagg acaaaatggt gtgggaaatc ctggtgtctg agcgggacat 600

cagagccctt atcccactgg ccaaggctga ggagctggtg aatacagcac 650

cactgactgg agtgccccag catgtccccg tgcgccttgt cactgtggac 700

ggcgggggg ccttggtgga ggtgacagag catgtcggct gcgagtctgc 750

caacacacag gtcctgcagg tgtctgaggc ctgtgatgcc gtgttcgtgg 800

ctggcaagga gagccggggc gcccgggggg tgcgagtgga cttctggtgg 850

cgccggctcc gcgcctcgct gcggctgacc gtgtgggccc cgctgctacc 900

gctgcgtatc gagctcaccg acaccaccct cgagcaggtc cgcggctgga 950

gggtacctgg ccctgctgaa gggcctgcgg aacccgctgc agaggcgtca 1000

gatgaggccg agcggcgcgc ccgtggctgc cacctgcagt accagcgggc 1050

cggtgtgcgc ttcctcgccc ccttcgcggc ccacccgctg gacggcggcc 1100

gccgcctcac gcacctgctt ggccccgact ggctgctaga cgtgtcccac 1150

ctcgtggcgc cacacgcccg cgtgctggac tcgcgtgtag cctctctgga 1200

gggtggccgt gtcgtggtgg gccgggagcc cggtgtcacc tccattgagg 1250

tgcgttcccc actgtctgac tccatcctgg gggagcaggc gctggctgtg 1300

acggacgaca aggtctcagt gctggagctg agggtgcagc cagtgatggg 1350

catctcgctg accttgagcc ggggcactgc ccaccccggg gaggtcacag 1400

ctacgtgctg ggcacagtca gcccttcccg ccccaaagca ggaggtggcc 1450

ctctccctat ggctgtcctt ctctgatcac actgtggccc cagctgagct 1500

ctacgaccgc cgtgacctgg gactgtccgt ctcagccgag gagcctggtg 1550

ccatcctgcc agctgaggag cagggtgccc agctcggggt ggtggtgagt 1600

ggggcaggcg ccgaggggct gccgctgcat gtggctctgc acccgcccga 1650

gccctgccgc cggggccgcc accgtgtgcc tctggcctct ggcaccgcct 1700

ggctggggct gccccctgcc tccactccag cccctgctct cccatccagc 1750

cctgcttgga gcccaccagc cacagaagcc accatgggtg gtaaacggca 1800

ggtggcaggc agtgtcgggg gcaacacagg tgtgaggggc aagtttgagc 1850

gggcagagga ggaggccagg aaggaggaga ccgaagccag ggaggaggag 1900

gaggaagagg aggaggagat ggtccctgcc cctcagcatg tcactgagct 1950

agagctgggc atgtacgccc tgctgggagt cttctgcgtg gccatcttca 2000

tcttcttggt caatggtgtg gtcttcgtcc tgcgctatca gcgcaaagaa 2050

cctcccgaca gtgccactga ccccacctcc ccccagcccc acaactgggt 2100

ctggctgggc actgaccagg aggaactgag ccgccagctg gaccggcagt 2150

cccctggccc gcccaagggg gaggggagct gcccctgtga gagtggggga 2200

ggaggggagg cccctaccct ggcccctggc cctcctgggg gcaccaccag 2250

ctcctcaagc accctggccc gaaaggaggc tgggggggcgg cggaagcgag 2300

tagagtttgt gacatttgtg ccagcccctc cagcccagtc acctgaggag 2350

cctgtagggg cccctgctgt gcagtccatc cttgtggcag gcgaggagga 2400

catccgctgg gtgtgtgagg acatggggct gaaggaccct gaggagcttc 2450

gcaactacat ggagaggatc cggggcagct cctgaccctc cacagccacc 2500

tggtcagcca ccagctgggg caacgagggt ggaggtccca ctgagcctct 2550

cgcctgcccc cgccactcgt ctggtgcttg ttgatccaag tcccctgcct 2600

ggtcccccac aaggactccc atccaggccc cctctgccct gcccccttgtc 2650

atggaccatg gtcgtgagga agggctcatg ccccttattt atgggaacca 2700

tctcattcta acagaataaa ccgagaagga aaccagaaaa aaaaaaaaaa 2750

aaaaa 2755

<210> 34
<211> 1795
<212> DNA
<213> Homo sapiens

<400> 34
acgcgtccgc ccacgcgtcc gcccacgcgt ccggtcgggg ccagagcgca 50

ggtgtacctg gcggccgtgc tggagcacct gaccgccgag atcctggagc 100

tggctggcaa cccggcccgc gacaagaaga cccgcatcat cctgcgccac 150

ctgtagctgg ccattcgcaa cggcgaggag cttaacaagc tgctgggcga 200

agtcaccatc gcgcagggcg gtgtcctgcc caacattcag ggcgtgcttc 250

tgccccagaa gaccaagagc caccacaagg ccaagggtga aaaccattca 300

ctaggagagg agaaacacaa tggccaccaa gacagagttg agtcccacag 350

caagggagag caagaacgca caagatatgc aagtggatga gacactgatc 400

cccaggaaag gtccaagttt atgttctgct cgctatggaa tagccctcgt 450

cttacatttc tgcaatttca caacgatagc acaaaatgtc atcatgaaca 500

tcaccatggt agccatggtc aacagcacaa gccctcaatc ccagctcaat 550

gattcctctg aggtgctgcc tgttgactca tttggtggcc taagtaaagc 600

cccaaagagt cttcctgcaa agtcctcaat acttgggggt cagtttgcaa 650

tttgggaaaa gtggggccct ccacaagaac gaagcagact ctgcagcatt 700

gctttatcag gaatgttact gggatgcttt actgccatcc tcataggtgg 750

cttcattagt gaaacccttg ggtggccctt tgtcttctat atctttggag 800

gtgttggctg tgtctgctgc cttctctggt ttgttgtgat ttatgatgac 850

ccctttttcct atccatggat aagcacctca gaaaaagaat acatcatatc 900

ctccttgaaa caacaggtcg ggtcttctaa gcagcctctt cccatcaaag 950

ctatgctcag atctctaccc atttggtcca tatgtttagg ctgtttcagc 1000

catcaatggt tagttagcac aatggttgta tacataccaa cttacatcag 1050

ctctgtgtac catgttaaca tcagagacaa tggacttcta tctgcccttc 1100

cttttattgt tgcctgggtc ataggcatgg tgggaggcta tctggcagat 1150

ttccttctaa ccaaaaagtt tagactcatc actgtgagga aaattgccac 1200

aattttagga agtctcccct cttcagcact cattgtgtct ctgccttacc 1250

tcaattccgg ctatatcaca gcaactgcct tgctgacgct ctcttgcgga 1300

ttaagcacat tgtgtcagtc agggatttat atcaatgtct tagatattgc 1350

tccaaggtat tccagttttc tcatgggagc atcaagagga ttttcgagca 1400

tagcacctgt cattgtaccc actgtcagcg gatttcttct tagtcaggac 1450

cctgagtttg ggtggaggaa tgtcttcttc ttgctgtttg ccgttaacct 1500

gttaggacta ctcttctacc tcatatttgg agaagcagat gtccaagaat 1550

gggctaaaga gagaaaactc actcgtttat gaagttatcc caccttggat 1600

ggaaaagtca ttaggcaccg tattgcataa aatagaaggc ttccgtgatg 1650

aaaataccag tgaaaagatt tttttttcct gtggctcttt tcaattatga 1700

gatcagttca ttattttatt cagacttttt tttgagagaa atgtaagatg 1750

aataaaaatt caaataaaat gataactaag aaaaaaaaaa aaaaa 1795

<210> 35
<211> 1795
<212> DNA
<213> Homo sapiens

<400> 35
acgcgtccgc ccacgcgtcc gcccacgcgt ccggtcgggg ccagagcgca 50

ggtgtacctg gcggccgtgc tggagcacct gaccgccgag atcctggagc 100

tggctggcaa cccggcccgc gacaagaaga cccgcatcat cctgcgccac 150

ctgtagctgg ccattcgcaa cggcgaggag cttaacaagc tgctgggcga 200

agtcaccatc gcgcagggcg gtgtcctgcc caacattcag ggcgtgcttc 250

tgccccagaa gaccaagagc caccacaagg ccaagggtga aaaccattca 300

ctaggagagg agaaacacaa tggccaccaa gacagagttg agtcccacag 350

caagggagag caagaacgca caagatatgc aagtggatga gacactgatc 400

cccaggaaag gtccaagttt atgttctgct cgctatggaa tagccctcgt 450

cttacatttc tgcaatttca caacgatagc acaaaatgtc atcatgaaca 500

tcaccatggt agccatggtc aacagcacaa gccctcaatc ccagctcaat 550

gattcctctg aggtgctgcc tgttgactca tttggtggcc taagtaaagc 600

cccaaagagt cttcctgcaa agtcctcaat acttgggggt cagtttgcaa 650

tttgggaaaa gtggggccct ccacaagaac gaagcagact ctgcagcatt 700

gctttatcag gaatgttact gggatgcttt actgccatcc tcataggtgg 750

cttcattagt gaaacccttg ggtggccctt tgtcttctat atctttggag 800

gtgttggctg tgtctgctgc cttctctggt ttgttgtgat ttatgatgac 850

ccctttttcct atccatggat aagcacctca gaaaaagaat acatcatatc 900

ctccttgaaa caacaggtcg ggtcttctaa gcagcctctt cccatcaaag 950

ctatgctcag atctctaccc atttggtcca tatgtttagg ctgtttcagc 1000

catcaatggt tagttagcac aatggttgta tacataccaa cttacatcag 1050

ctctgtgtac catgttaaca tcagagacaa tggacttcta tctgcccttc 1100

cttttattgt tgcctgggtc ataggcatgg tgggaggcta tctggcagat 1150

ttccttctaa ccaaaaagtt tagactcatc actgtgagga aaattgccac 1200

aattttagga agtctcccct cttcagcact cattgtgtct ctgccttacc 1250

tcaattccgg ctatatcaca gcaactgcct tgctgacgct ctcttgcgga 1300

ttaagcacat tgtgtcagtc agggatttat atcaatgtct tagatattgc 1350

tccaaggtat tccagttttc tcatgggagc atcaagagga ttttcgagca 1400

tagcacctgt cattgtaccc actgtcagcg gatttcttct tagtcaggac 1450

cctgagtttg ggtggaggaa tgtcttcttc ttgctgtttg ccgttaacct 1500

gttaggacta ctcttctacc tcatatttgg agaagcagat gtccaagaat 1550

gggctaaaga gagaaaactc actcgtttat gaagttatcc caccttggat 1600

ggaaaagtca ttaggcaccg tattgcataa aatagaaggc ttccgtgatg 1650

aaaataccag tgaaaagatt tttttttcct gtggctcttt tcaattatga 1700

gatcagttca ttattttatt cagacttttt tttgagagaa atgtaagatg 1750

aataaaaatt caaataaaat gataactaag aaaaaaaaaa aaaaa 1795

<210> 36
<211> 4524

<212> DNA
<213> Homo sapiens

<400> 36
gctcgctggg ccgctgctcc cgggtgtccc aggcccggcc ggtgcgcaga 50

gcatggcggg tgcgggcccg aagcggcgcg cgctagcggc gccggcggcc 100

gaggagaagg aagaggcgcg ggagaagatg ctggccgcca agagcgcgga 150

cggctcggcg ccggcaggcg agggcgaggg cgtgaccctg cagcggaaca 200

tcacgctgct caacggcgtg gccatcatcg tggggaccat tatcggctcg 250

ggcatcttcg tgacgcccac gggcgtgctc aaggaggcag gctcgccggg 300

gctggcgctg gtggtgtggg ccgcgtgcgg cgtcttctcc atcgtgggcg 350

cgctctgcta cgcggagctc ggcaccacca tctccaaatc gggcggcgac 400

tacgcctaca tgctggaggt ctacggctcg ctgcccgcct tcctcaagct 450

ctggatcgag ctgctcatca tccggccttc atcgcagtac atcgtggccc 500

tggtcttcgc cacctacctg ctcaagccgc tcttccccac ctgcccggtg 550

cccgaggagg cagccaagct cgtggcctgc ctgtgcgtgc tgctgctcac 600

ggccgtgaac tgctacagcg tgaaggccgc cacccgggtc caggatgcct 650

ttgccgccgc caagctcctg gccctggccc tgatcatcct gctgggcttc 700

gtccagatcg gaaagggtga tgtgtccaat ctagatccca agttctcatt 750

tgaaggcacc aaactggatg tggggaacat tgtgctggca ttatacagcg 800

gcctctttgc ctatggagga tggaattact tgaatttcgt cacagaggaa 850

atgatcaacc cctacagaaa cctgcccctg gccatcatca tctccctgcc 900

catcgtgacg ctggtgtacg tgctgaccaa cctggcctac ttcaccaccc 950

tgtccaccga gcagatgctg tcgtccgagg ccgtggccgt ggacttcggg 1000

aactatcacc tgggcgtcat gtcctggatc atccccgtct tcgtgggcct 1050

gtcctgcttt ggctccgtca atgggtccct gttcacatcc tccaggctct 1100

tcttcgtggg gtcccgggaa ggccacctgc cctccatcct ctccatgatc 1150

cacccacagc tcctcacccc cgtgccgtcc ctcgtgttca cgtgtgtgat 1200

gacgctgctc tacgccttct ccaaggacat cttctccgtc atcaacttct 1250

tcagcttctt caactggctc tgcgtggccc tggccatcat cggcatgatc 1300

tggctgcgcc acagaaagcc tgagcttgag cggcccatca aggtgaacct 1350

ggccctgcct gtgttcttca tcctggcctg cctcttcctg atcgccgtct 1400

ccttctggaa gacacccgtg gagtgtggca tcggcttcac catcatcctc 1450

agcgggctgc ccgtctactt cttcggggtc tggtggaaaa acaagcccaa 1500

gtggctcctc cagggcatct tctccacgac cgtcctgtgt cagaagctca 1550

tgcaggtggt cccccaggag acatagccag gaggccgagt ggctgccgga 1600

ggagcatgcg cagaggccag ttaaagtaga tcacctcctc gaacccactc 1650

cggttccccg caacccacag ctcagctgcc catcccagtc ctcgccgtcc 1700

ctcccaggtc gggcagtgga ggctgctgtg aaaactctgg tacgaatctc 1750

atccctcaac tgagggccag ggacccaggt gtgcctgtgc tcctgcccag 1800

gagcagcttt tggtctcctt gggccctttt tcccttccct cctttgttta 1850

cttatatata tatttttttt aaacttaaat tttgggtcaa cttgacacca 1900

ctaagatgat tttttaagga gctgggggaa ggcaggagcc ttcctttctc 1950

ctgccccaag ggcccagacc ctgggcaaac agagctactg agacttggaa 2000

cctcattgct accacagact tgcactgaag ccagacagct gcccagacac 2050

atgggcttgt gacattcgtg aaaaccaacc ctgtgggctt atgtctctgc 2100

cttagggttt gcagagtgga aactcagccg tagggtggca ctgggagggg 2150

gtgggggatc tgggcaaggt gggtgattcc tcccaggagg tgcttgaggc 2200

cccgatggac tcctgaccat aatcctagcc ccgagacacc atcctgagcc 2250

agggaacagc cccagggttg gggggtgccg gcatctcccc tagctcacca 2300

ggcctggcct ctgggcagtg tggcctcttg gctatttctg ttccagtttt 2350

ggaggctgag ttctggttca tgcagacaaa gccctgtcct tcagtcttct 2400

agaaacagag acaagaaagg cagacacacc gcggccaggc acccatgtgg 2450

gcgcccaccc tgggctccac acagcagtgt cccctgcccc agaggtcgca 2500

gctaccctca gcctccaatg cattggcctc tgtaccgccc ggcagcccct 2550

tctggccggt gctgggttcc cactcccggc ctaggcacct ccccgctctc 2600

cctgtcacgc tcatgtcctg tcctggtcct gatgcccgtt gtctaggaga 2650

cagagccaag cactgctcac gtctctgccg cctgcgtttg gaggcccctg 2700

ggctctcacc cagtccccac ccgcctgcag agagggaact agggcacccc 2750

ttgtttctgt tgttcccgtg aatttttttc gctatgggag gcagccgagg 2800

cctggccaat gcggcccact ttcctgagct gtcgctgcct ccatggcagc 2850

agccaaggac ccccagaaca agaagacccc cccgcaggat ccctcctgag 2900

ctcgggggggc tctgccttct caggccccgg gcttcccttc tccccagcca 2950

gaggtggagc caagtggtcc agcgtcactc cagtgctcag ctgtggctgg 3000

aggagctggc ctgtggcaca gccctgagtg tcccaagccg ggagccaacg 3050

aagccggaca cggcttcact gaccagcggc tgctcaagcc gcaagctctc 3100

agcaagtgcc cagtggagcc tgccgccccc acctgggcac cgggaccccc 3150

tcaccatcca gtgggcccgg agaaacctga tgaacagttt ggggactcag 3200

gaccagatgt ccgtctctct tgcttgagga atgaagacct ttattcaccc 3250

ctgccccgtt gcttcccgct gcacatggac agacttcaca gcgtctgctc 3300

ataggacctg catccttcct ggggacgaat tccactcgtc caagggacag 3350

cccacggtct ggaggccgag gaccaccagc aggcaggtgg actgactgtg 3400

ttgggcaaga cctcttccct ctgggcctgt tctcttggct gcaaataagg 3450

acagcagctg gtgccccacc tgcctggtgc attgctgtgt gaatccagga 3500

ggcagtggac atcgtaggca gccacggccc caggtccagg agaagtgctc 3550

cctggaggca cggaccactg cttcccactg gggccggcgg ggcccacgca 3600

cgacgtcagc ctcttacctt cccgcctcgg ctaggggtcc tcgggatgcc 3650

gttctgttcc aacctcctgt tctgggaggt ggacatgcct caaggataca 3700

gggagccggc ggcctctcga cggcacgcac ttcctgttgg ctgctgcggc 3750

tgtgggcgag catgggggct gccagcgtct gttgtggaaa gtagctgcta 3800

gtgaaatggc tggggccgct ggggtccgtc ttcacactgc gcaggtctct 3850

tctgggcgtc tgagctgggg tgggagctcc tccgcagaag gttggtgggg 3900

ggtccagtct gtgatccttg gtgctgtgtg ccccactcca gcctggggac 3950

cccacttcag aaggtagggg ccgtgtcccg cggtgctgac tgaggcctgc 4000

ttccccctcc ccctcctgct gtgctggaat tccacaggga ccagggccac 4050

cgcaggggac tgtctcagaa gacttgattt ttccgtccct ttttctccac 4100

actccactga caaacgtccc cagcggtttc cacttgtggg cttcaggtgt 4150

tttcaagcac aacccaccac aacaagcaag tgcattttca gtcgttgtgc 4200

ttttttgttt tgtgctaacg tcttactaat ttaaagatgc tgtcggcacc 4250

atgtttattt atttccagtg gtcatgctca gccttgctgc tctgcgtggc 4300

gcaggtgcca tgcctgctcc ctgtctgtgt cccagccacg cagggccatc 4350

cactgtgacg tcggccgacc aggctggaca ccctctgccg agtaatgacg 4400

tgtgtggctg ggaccttctt tattctgtgt taatggctaa cctgttacac 4450

tgggctgggt tgggtagggt gttctggctt ttttgtgggg tttttatttt 4500

taaagaaaca ctcaatcatc ctag 4524

<210> 37
<211> 1590
<212> DNA
<213> Homo sapiens

<400> 37
attcgccctt tgctcccggg tgtcccaggc ccggccggtg cgcagagcat 50

ggcgggtgcg ggcccgaagc ggcgcgcgct agcggcgccg gcggccgagg 100

agaaggaaga ggcgcgggag aagatgctgg ccgccaagag cgcggacggc 150

tcggcgccgg caggcgaggg cgagggcgtg accctgcagc ggaacatcac 200

gctgctcaac ggcgtggcca tcatcgtggg gaccattatc ggctcgggca 250

tcttcgtgac gcccacgggc gtgctcaagg aggcaggctc gccggggctg 300

gcgctggtgg tgtgggccgc gtgcggcgtc ttctccatcg tgggcgcgct 350

ctgctacgcg gagctcggca ccaccatctc caaatcgggc ggcgactacg 400

cctacatgct ggaggtctac ggctcgctgc ccgccttcct caagctctgg 450

atcgagctgc tcatcatccg gccttcatcg cagtacatcg tggccctggt 500

cttcgccacc tacctgctca agccgctctt ccccacctgc ccggtgcccg 550

aggaggcagc caagctcgtg gcctgcctct gcgtgctgct gctcacggcc 600

gtgaactgct acagcgtgaa ggccgccacc cgggtccagg atgcctttgc 650

cgccgccaag ctcctggccc tggccctgat catcctgctg ggcttcgtcc 700

agatcgggaa gggtgatgtg tccaatctag atcccaactt ctcatttgaa 750

ggcaccaaac tggatgtggg gaacattgtg ctggcattat acagcggcct 800

ctttgcctat ggaggatgga attacttgaa tttcgtcaca gaggaaatga 850

tcaaccccta cagaaacctg cccctggcca tcatcatctc cctgcccatc 900

gtgacgctgg tgtacgtgct gaccaacctg gcctacttca ccaccctgtc 950

caccgagcag atgctgtcgt ccgaggccgt ggccgtggac ttcgggaact 1000

atcacctggg cgtcatgtcc tggatcatcc ccgtcttcgt gggcctgtcc 1050

tgcttcggct ccgtcaatgg gtccctgttc acatcctcca ggctcttctt 1100

cgtggggtcc cgggaaggcc acctgccctc catcctctcc atgatccacc 1150

cacagctcct cacccccgtg ccgtccctcg tgttcacgtg tgtgatgacg 1200

ctgctctacg ccttctccaa ggacatcttc tccgtcatca acttcttcag 1250

cttcttcaac tggctctgcg tggccctggc catcatcggc atgatctggc 1300

tgcgccacag aaagcctgag cttgagcggc ccatcaaggt gaacctggcc 1350

ctgcctgtgt tcttcatcct ggcctgcctc ttcctgatcg ccgtctcctt 1400

ctggaagaca cccgtggagt gtggcatcgg cttcaccatc atcctcagcg 1450

ggctgcccgt ctacttcttc ggggtctggt ggaaaaacaa gcccaagtgg 1500

ctcctccagg gcatcttctc cacgaccgtc ctgtgtcaga agctcatgca 1550

ggtggtcccc caggagacat agccaggagg ccgaagggcg 1590

<210> 38
<211> 2436
<212> DNA
<213> Homo sapiens

<400> 38
ccaggctctc cacccccact tcccaattga ggaaaccgag gcagaggagg 50

ctcaggtgtg gccaatcacc ctgcacatca gagttaccct gggcagggcc 100

cactgagacc tgggaggggc cactcgggac ctggagggct gggggctgcc 150

cgggcgttag gggtaaagct ccctacccaa ctgcgcagaa ggcctcagag 200

gcctgggggc tgggcttccc ctttcacatc gccctttaga ggcccacgtg 250

tgggcattgg cccgcgatct gaaaggggct gtcctgttcc tcatgggcgc 300

tgccagcgcc acgcactcct ctttctgcct ggccggccac tcccgtctgc 350

tgtgacgcgc ggacagagag ctaccggtgg acccacggtg cctccctccc 400

tgggatctac acagaccatg gccttgccaa cggctcgacc cctgttgggg 450

tcctgtggga cccccgccct cggcagcctc ctgttcctgc tcttcagcct 500

cggatgggtg cagccctcga ggaccctggc tggagagaca gggcaggagg 550

ctgcacccct ggacggagtc ctggccaacc cacctaacat ttccagcctc 600

tcccctcgcc aactccttgg cttcccgtgt gcggaggtgt ccggcctgag 650

cacggagcgt gtccgggagc tggctgtggc cttggcacag aagaatgtca 700

agctctcaac agagcagctg cgctgtctgg ctcaccggct ctctgagccc 750

cccgaggacc tggacgccct cccattggac ctgctgctat tcctcaaccc 800

agatgcgttc tcggggcccc aggcctgcac ccgtttcttc tcccgcatca 850

cgaaggccaa tgtggacctg ctcccgaggg gggctcccga gcgacagcgg 900

ctgctgcctg cggctctggc ctgctggggt gtgcgggggt ctctgctgag 950

cgaggctgat gtgcgggctc tgggaggcct ggcttgcgac ctgcctgggc 1000

gctttgtggc cgagtcggcc gaagtgctgc taccccggct ggtgagctgc 1050

ccgggacccc tggaccagga ccagcaggag gcagccaggg cggctctgca 1100

gggcggggga ccccccctacg gccccccgtc gacatggtct gtctccacga 1150

tggacgctct gcggggcctg ctgcccgtgc tgggccagcc catcatccgc 1200

agcatcccgc agggcatcgt ggccgcgtgg cggcaacgct cctctcggga 1250

cccatcctgg cggcagcctg aacggaccat cctccggccg cggttccggc 1300

gggaagtgga gaagacagcc tgtccttcag gcaagaaggc ccgcgagata 1350

gacgagagcc tcatcttcta caagaagtgg gagctggaag cctgcgtgga 1400

tgcggccctg ctggccaccc agatggaccg cgtgaacgcc atccccttca 1450

cctacgagca gctggacgtc ctaaagcata aactggatga gctctaccca 1500

caaggttacc ccgagtctgt gatccagcac ctgggctacc tcttcctcaa 1550

gatgagccct gaggacattc gcaagtggaa tgtgacgtcc ctggagaccc 1600

tgaaggcttt gcttgaagtc aacaaagggc acgaaatgag tcctcaggtg 1650

gccaccctga tcgaccgctt tgtgaaggga aggggccagc tagacaaaga 1700

caccctagac accctgaccg ccttctaccc tgggtacctg tgctccctca 1750

gccccgagga gctgagctcc gtgcccccca gcagcatctg ggcggtcagg 1800

ccccaggacc tggacacgtg tgacccaagg cagctggacg tcctctatcc 1850

caaggcccgc cttgctttcc agaacatgaa cgggtccgaa tacttcgtga 1900

agatccagtc cttcctgggt ggggccccca cggaggattt gaaggcgctc 1950

agtcagcaga atgtgagcat ggacttggcc acgttcatga agctgcggac 2000

ggatgcggtg ctgccgttga ctgtggctga ggtgcagaaa cttctgggac 2050

cccacgtgga gggcctgaag gcggaggagc ggcaccgccc ggtgcgggac 2100

tggatcctac ggcagcggca ggacgacctg gacacgctgg ggctggggct 2150

acagggcggc atccccaacg gctacctggt cctagacctc agcgtgcaag 2200

aggccctctc ggggacgccc tgcctcctag gacctggacc tgttctcacc 2250

gtcctggcac tgctcctagc ctccaccctg gcctgagggc cccactccct 2300

tgctggcccc agccctgctg gggatccccg cctggccagg agcaggcacg 2350

ggtgatcccc gttccacccc aagagaactc gcgctcagta aacgggaaca 2400

tgccccctgc agacacgtaa aaaaaaaaaa aaaaaa 2436

<210> 39
<211> 1979
<212> DNA
<213> Homo sapiens

<400> 39

tgggatctac acagaccatg gccttgccaa cggctcgacc cctgttgggg 50

tcctgtggga cccccgccct cggcagcctc ctgttcctgc tcttcagcct 100

cggatgggtg cagccctcga ggaccctggc tggagagaca gggcaggagg 150

ctgcgcccct ggacggagtc ctggccaacc cacctaacat ttccagcctc 200

tcccctcgcc aactccttgg cttcccgtgt gcggaggtgt ccggcctgag 250

cacggagcgt gtccgggagc tggctgtggc cttggcacag aagaatgtca 300

agctctcaac agagcagctg cgctgtctgg ctcaccggct ctctgagccc 350

cccgaggacc tggacgccct cccattggac ctgctgctat tcctcaaccc 400

agatgcgttc tcggggcccc aggcctgcac ccgtttcttc tcccgcatca 450

cgaaggccaa tgtggacctg ctcccgaggg gggctcccga gcgacagcgg 500

ctgctgcctg cggctctggc ctgctggggt gtgcgggggt ctctgctgag 550

cgaggctgat gtgcgggctc tgggaggcct ggcttgcgac ctgcctgggc 600

gctttgtggc cgagtcggcc gaagtgctgc taccccggct ggtgagctgc 650

ccgggacccc tggaccagga ccagcaggag gcagccaggg cggctctgca 700

gggcggggga cccccctacg gccccccgtc gacatggtct gtctccacga 750

tggacgctct gcggggcctg ctgcccgtgc tgggccagcc catcatccgc 800

agcatcccgc agggcatcgt ggccgcgtgg cggcaacgct cctctcggga 850

cccatcctgg cggcagcctg aacggaccat cctccggccg cggttccggc 900

gggaagtgga gaagacagcc tgtccttcag gcaagaaggc ccgcgagata 950

gacgagagcc tcatcttcta caagaagtgg gagctggaag cctgcgtgga 1000

tgcggccctg ctggccaccc agatggaccg cgtgaacgcc atccccttca 1050

cctacgagca gctggacgtc ctaaagcata aactggatga gctctaccca 1100

caaggttacc ccgagtctgt gatccagcac ctgggctacc tcttcctcaa 1150

gatgagccct gaggacattc gcaagtggaa tgtgacgtcc ctggagaccc 1200

tgaaggcttt gcttgaagtc aacaaagggc acgaaatgag tcctcaggtg 1250

gccaccctga tcgaccgctt tgtgaaggga aggggccagc tagacaaaga 1300

caccctagac accctgaccg ccttctaccc tgggtacctg tgctccctca 1350

gccccgagga gctgagctcc gtgccccccа gcagcatctg ggcggtcagg 1400

ccccaggacc tggacacgtg tgacccaagg cagctggacg tcctctatcc 1450

caaggcccgc cttgctttcc agaacatgaa cgggtccgaa tacttcgtga 1500

agatccagtc cttcctgggt ggggcccccа cggaggattt gaaggcgctc 1550

agtcagcaga atgtgagcat ggacttggcc acgttcatga agctgcggac 1600

ggatgcggtg ctgccgttga ctgtggctga ggtgcagaaa cttctgggac 1650

cccacgtgga gggcctgaag gcggaggagc ggcaccgccc ggtgcgggac 1700

tggatcctac ggcagcggca ggacgacctg gacacgctgg ggctggggct 1750

acagggcggc atccccaacg gctacctggt cctagacctc agcgtgcaag 1800

aggccctctc ggggacgccc tgcctcctag gacctggacc tgttctcacc 1850

gtcctggcac tgctcctagc ctccaccctg gcctgagggc cccactccct 1900

tgctggcccc agccctgctg gggatccccg cctggccagg agcaggcacg 1950

ggtgatcccc gttccacccc aagagaact 1979

<210> 40
<211> 2456
<212> DNA
<213> Homo sapiens

<400> 40
ataacagcat gaagtgccgt ggaactggaa taggcgtgtc ctctccctcg 50

accctccccc tccttgtccc tctgctcacc cctcgctcgt tccctccctc 100

cggcgagggc cgcctttata acaactgctc agagtgcgag ggcgggatag 150

ctgtccaagg tctcccccag cactgaggag ctcgcctgct gccctcttgc 200

gcgcgggaag cagcaccaag ttcacggcca acgccttggc actagggtcc 250

agaatggcta caacagtccc tgatggttgc cgcaatggcc tgaaatccaa 300

gtactacaga ctttgtgata aggctgaagc ttggggcatc gtcctagaaa 350

cggtggccac agccggggtt gtgacctcgg tggccttcat gctcactctc 400

ccgatcctcg tctgcaaggt gcaggactcc aacaggcgaa aaatgctgcc 450

tactcagttt ctcttcctcc tgggtgtgtt gggcatcttt ggcctcacct 500

tcgccttcat catcggactg gacgggagca cagggcccac acgcttcttc 550

ctctttggga tcctcttttc catctgcttc tcctgcctgc tggctcatgc 600

tgtcagtctg accaagctcg tccgggggag gaagcccctt tccctgttgg 650

tgattctggg tctggccgtg ggcttcagcc tagtccagga tgttatcgct 700

attgaatata ttgtcctgac catgaatagg accaacgtca atgtcttttc 750

tgagctttcc gctcctcgtc gcaatgaaga ctttgtcctc ctgctcacct 800

acgtcctctt cttgatggcg ctgaccttcc tcatgtcctc cttcaccttc 850

tgtggttcct tcacgggctg gaagagacat ggggcccaca tctacctcac 900

gatgctcctc tccattgcca tctgggtggc ctggatcacc ctgctcatgc 950

ttcctgactt tgaccgcagg tgggatgaca ccatcctcag ctccgccttg 1000

gctgccaatg gctgggtgtt cctgttggct tatgttagtc ccgagttttg 1050

gctgctcaca aagcaacgaa accccatgga ttatcctgtt gaggatgctt 1100

tctgtaaacc tcaactcgtg aagaagagct atggtgtgga gaacagagcc 1150

tactctcaag aggaaatcac tcaaggtttt gaagagacag gggacacgct 1200

ctatgccccc tattccacac attttcagct gcagaaccag cctccccaaa 1250

aggaattctc catcccacgg gcccacgctt ggccgagccc ttacaaagac 1300

tatgaagtaa agaaagaggg cagctaactc tgtcctgaag agtgggacaa 1350

atgcagccgg gcggcagatc tagcgggagc tcaaagggat gtgggcgaaa 1400

tcttgagtct tctgagaaaa ctgtacaaga cactacggga acagtttgcc 1450

tccctcccag cctcaaccac aattcttcca tgctggggct gatgtgggct 1500

agtaagactc cagttcttag aggcgctgta gtatttttt ttttttgtct 1550

catcctttgg atacttcttt taagtgggag tctcaggcaa ctcaagttta 1600

gacccttact ctttttgttt gtttttttgaa acaggatctt gctctgtcac 1650

ccaggcttga gtgcagtggt gcgatcacag cccagtgcag cctcgaccac 1700

ctgtgctcaa gcaatcctcc catctccatc tcccaaagtg ctgggatgac 1750

aggcgtgagc cacagctccc agcctaggcc cttaatcttg ctgttatttt 1800

ccatggacta aaggtctggt catctgagct cacgctggct cacacagctc 1850

taggggcctg ctcctctaac tcacagtggg ttttgtgagg ctctgtggcc 1900

cagagcagac ctgcatatct gagcaaaaat agcaaaagcc tctctcagcc 1950

cactggcctg aatctacact ggaagccaac ttgctggcac ccccgctccc 2000

caacccttct tgcctgggta ggagaggcta aagatcaccc taaatttact 2050

catctctcta gtgctgcctc acattgggcc tcagcagctc cccagcacca 2100

attcacaggt cacccctctc ttcttgcact gtccccaaac ttgctgtcaa 2150

ttccgagatc taatctcccc ctacgctctg ccaggaattc tttcagacct 2200

cactagcaca agcccggttg ctccttgtca ggagaatttg tagatcattc 2250

tcacttcaaa ttcctggggc tgatacttct ctcatcttgc accccaacct 2300

ctgtaaatag atttaccgca tttacggctg cattctgtaa gtgggcatgg 2350

tctcctaatg gaggagtgtt cattgtataa taagttattc acctgagtat 2400

gcaataaaga tgtggtggcc actctttcat ggtggtggca gcaaaaaaaa 2450

aaaaaa 2456

<210> 41
<211> 3213
<212> DNA
<213> Homo sapiens

<400> 41
agagactcaa gatgattccc tttttaccca tgttttctct actattgctg 50

cttattgtta accctataaa cgccaacaat cattatgaca agatcttggc 100

tcatagtcgt atcaggggtc gggaccaagg cccaaatgtc tgtgcccttc 150

aacagatttt gggcaccaaa aagaaatact tcagcacttg taagaactgg 200

tataaaaagt ccatctgtgg acagaaaacg actgttttat atgaatgttg 250

ccctggttat atgagaatgg aaggaatgaa aggctgccca gcagttttgc 300

ccattgacca tgtttatggc actctgggca tcgtgggagc caccacaacg 350

cagcgctatt ctgacgcctc aaaactgagg gaggagatcg agggaaaggg 400

atccttcact tactttgcac cgagtaatga ggcttgggac aacttggatt 450

ctgatatccg tagaggtttg gagagcaacg tgaatgttga attactgaat 500

gctttacata gtcacatgat taataagaga atgttgacca aggacttaaa 550

aaatggcatg attattcctt caatgtataa caatttgggg cttttcatta 600

accattatcc taatggggtt gtcactgtta attgtgctcg aatcatccat 650

gggaaccaga ttgcaacaaa tggtgttgtc catgtcattg accgtgtgct 700

tacacaaatt ggtacctcaa ttcaagactt cattgaagca gaagatgacc 750

tttcatcttt tagagcagct gccatcacat cggacatatt ggaggccctt 800

ggaagagacg gtcacttcac actctttgct cccaccaatg aggcttttga 850

gaaacttcca cgaggtgtcc tagaaaggtt catgggagac aaagtggctt 900

ccgaagctct tatgaagtac cacatcttaa atactctcca gtgttctgag 950

tctattatgg gaggagcagt ctttgagacg ctggaaggaa atacaattga 1000

gataggatgt gacggtgaca gtataacagt aaatggaatc aaaatggtga 1050

acaaaaagga tattgtgaca aataatggtg tgatccattt gattgatcag 1100

gtcctaattc ctgattctgc caaacaagtt attgagctgg ctggaaaaca 1150

gcaaaccacc ttcacggatc ttgtggccca attaggcttg gcatctgctc 1200

tgaggccaga tggagaatac actttgctgg cacctgtgaa taatgcattt 1250

tctgatgata ctctcagcat ggttcagcgc ctccttaaat taattctgca 1300

gaatcacata ttgaaagtaa aagttggcct taatgagctt tacaacgggc 1350

aaatactgga aaccatcgga ggcaaacagc tcagagtctt cgtatatcgt 1400

acagctgtct gcattgaaaa ttcatgcatg gagaaaggga gtaagcaagg 1450

gagaaacggt gcgattcaca tattccgcga gatcatcaag ccagcagaga 1500

aatccctcca tgaaaagtta aaacaagata agcgctttag caccttcctc 1550

agcctacttg aagctgcaga cttgaaagag ctcctgacac aacctggaga 1600

ctggacatta tttgtgccaa ccaatgatgc ttttaaggga atgactagtg 1650

aagaaaaaga aattctgata cgggacaaaa atgctcttca aaacatcatt 1700

ctttatcacc tgacaccagg agttttcatt ggaaaaggat ttgaacctgg 1750

tgttactaac attttaaaga ccacacaagg aagcaaaatc tttctgaaag 1800

aagtaaatga tacacttctg gtgaatgaat tgaaatcaaa agaatctgac 1850

atcatgacaa caaatggtgt aattcatgtt gtagataaac tcctctatcc 1900

agcagacaca cctgttggaa atgatcaact gctggaaata cttaataaat 1950

taatcaaata catccaaatt aagtttgttc gtggtagcac cttcaaagaa 2000

atccccgtga ctgtctatac aactaaaatt ataaccaaag ttgtggaacc 2050

aaaaattaaa gtgattgaag gcagtcttca gcctattatc aaaactgaag 2100

gacccacact aacaaaagtc aaaattgaag gtgaacctga attcagactg 2150

attaaagaag gtgaaacaat aactgaagtg atccatggag agccaattat 2200

taaaaaatac accaaaatca ttgatggagt gcctgtggaa ataactgaaa 2250

aagagacacg agaagaacga atcattacag gtcctgaaat aaaatacact 2300

aggatttcta ctggaggtgg agaaacagaa gaaactctga agaaattgtt 2350

acaagaagag gtcaccaagg tcaccaaatt cattgaaggt ggtgatggtc 2400

atttatttga agatgaagaa attaaaagac tgcttcaggg agacacaccc 2450

gtgaggaagt tgcaagccaa caaaaaagtt caaggttcta gaagacgatt 2500

aagggaaggt cgttctcagt gaaaatccaa aaaccagaaa aaaatgttta 2550

tacaacccta agtcaataac ctgaccttag aaaattgtga gagccaagtt 2600

gacttcagga actgaaacat cagcacaaag aagcaatcat caaataattc 2650

tgaacacaaa tttaatattt ttttttctga atgagaaaca tgagggaaat 2700

tgtggagtta gcctcctgtg gtaaaggaat tgaagaaaat ataacacctt 2750

acacccтttt tcatcttgac attaaaagtt ctggctaact ttggaatcca 2800

ttagagaaaa atccttgtca ccagattcat tacaattcaa atcgaagagt 2850

tgtgaactgt tatcccattg aaaagaccga gccttgtatg tatgttatgg 2900

atacataaaa tgcacgcaag ccattatctc tccatgggaa gctaagttat 2950

aaaaataggt gcttggtgta caaaactttt tatatcaaaa ggctttgcac 3000

atttctatat gagtgggttt actggtaaat tatgttattt tttacaacta 3050

attttgtact ctcagaatgt ttgtcatatg cttcttgcaa tgcatatttt 3100

ttaatctcaa acgtttcaat aaaaccattt ttcagatata aagagaatta 3150

cttcaaattg agtaattcag aaaaactcaa gatttaagtt aaaaagtggt 3200

ttggacttgg gaa 3213

<210> 42
<211> 2450
<212> DNA
<213> Homo sapiens

<400> 42
cccttgagac tcaagatgat tccctttta cccatgtttt ctctactatt 50

gctgcttatt gttaacccta taaacgccaa caatcattat gacaagatct 100

tggctcatag tcgtatcagg ggtcgggacc aaggcccaaa tgtctgtgcc 150

cttcaacaga ttttgggcac caaaaagaaa tacttcagca cttgtaagaa 200

ctggtataaa aagtccatct gtggacagaa aacgactgtg ttatatgaat 250

gttgccctgg ttatatgaga atggaaggaa tgaaaggctg cccagcagtt 300

ttgcccattg accatgttta tggcactctg ggcatcgtgg gagccaccac 350

aacgcagcgc tattctgacg cctcaaaact gagggaggag atcgagggaa 400

agggatcctt cacttacttt gcaccgagta atgaggcttg ggacaacttg 450

gattctgata tccgtagagg tttggagagc aacgtgaatg ttgaattact 500

gaatgcttta catagtcaca tgattaataa gagaatgttg accaaggact 550

taaaaaatgg catgattatt ccttcaatgt ataacaattt ggggcttttc 600

attaaccatt atcctaatgg ggttgtcact gttaattgtg ctcgaatcat 650

ccatgggaac cagattgcaa caaatggtgt tgtccatgtc attgaccgtg 700

tgcttacaca aattggtacc tcaattcaag acttcattga agcagaagat 750

gacctttcat cttttagagc agctgccatc acatcggaca tattggaggc 800

ccttggaaga gacggtcact tcacactctt tgctcccacc aatgaggctt 850

ttgagaaact tccacgaggt gtcctagaaa ggatcatggg agacaaagtg 900

gcttccgaag ctcttatgaa gtaccacatc ttaaatactc tccagtgttc 950

tgagtctatt atgggaggag cagtctttga gacgctggaa ggaaatacaa 1000

ttgagatagg atgtgacggt gacagtataa cagtaaatgg aatcaaaatg 1050

gtgaacaaaa aggatattgt gacaaataat ggtgtgatcc atttgattga 1100

tcaggtccta attcctgatt ctgccaaaca agttattgag ctggctggaa 1150

aacagcaaac caccttcacg gatcttgtgg cccaattagg cttggcatct 1200

gctctgaggc cagatggaga atacactttg ctggcacctg tgaataatgc 1250

attttctgat gatactctca gcatggatca gcgcctcctt aaattaattc 1300

tgcagaatca catattgaaa gtaaaagttg gccttaatga gctttacaac 1350

gggcaaatac tggaaaccat cggaggcaaa cagctcagag tcttcgtata 1400

tcgtacagct gtctgcattg aaaattcatg catggagaaa gggagtaagc 1450

aagggagaaa cggtgcgatt cacatattcc gcgagatcat caagccagca 1500

gagaaatccc tccatgaaaa gttaaaacaa gataagcgct ttagcacctt 1550

cctcagccta cttgaagctg cagacttgaa agagctcctg acacaacctg 1600

gagactggac attatttgtg ccaaccaatg atgcttttaa gggaatgact 1650

agtgaagaaa aagaaattct gatacgggac aaaaatgctc ttcaaaacat 1700

cattctttat cacctgacac caggagtttt cattggaaaa ggatttgaac 1750

ctggtgttac taacatttta aagaccacac aaggaagcaa aatctttctg 1800

aaagaagtaa atgatacact tctggtgaat gaattgaaat caaaagaatc 1850

tgacatcatg acaacaaatg gtgtaattca tgttgtagat aaactcctct 1900

atccagcaga cacacctgtt ggaaatgatc aactgctgga aatacttaat 1950

aaattaatca aatacatcca aattaagttt gttcgtggta gcaccttcaa 2000

agaaatcccc gtgactgtct atagacccac actaacaaaa gtcaaaattg 2050

aaggtgaacc tgaattcaga ctgattaaag aaggtgaaac aataactgaa 2100

gtgatccatg gagagccaat tattaaaaaa tacaccaaaa tcattgatgg 2150

agtgcctgtg gaaataactg aaaaagagac acgagaagaa cgaatcatta 2200

caggtcctga aataaaatac actaggattt ctactggagg tggagaaaca 2250

gaagaaactc tgaagaaatt gttacaagaa gaggtcacca aggtcaccaa 2300

attcattgaa ggtggtgatg gtcatttatt tgaagatgaa gaaattaaaa 2350

gactgcttca gggagacaca cccgtgagga agttgcaagc caacaaaaaa 2400

gttcaaggat ctagaagacg attaagggaa ggtcgttctc agtgaagggc 2450

<210> 43

<211> 2276
<212> DNA
<213> Homo sapiens

<400> 43
ccccttgagac tcaagatgat tccctttta cccatgtttt ctctactatt 50

gctgcttatt gttaacccta taaacgccaa caatcattat gacaagatct 100

tggctcatag tcgtatcagg ggtcgggacc aaggcccaaa tgtctgtgcc 150

cttcaacaga ttttgggcac caaaaagaaa tacttcagca cttgtaagaa 200

ctggtataaa aagtccatct gtggacagaa aacgactgtg ttatatgaat 250

gttgccctgg ttatatgaga atggaaggaa tgaaaggctg cccagcagtt 300

ttgcccattg accatgttta tggcactctg ggcatcgtgg gagccaccac 350

aacgcagcgc tattctgacg cctcaaaact gagggaggag atcgagggaa 400

agggatcctt cacttacttt gcaccgagta atgaggcttg ggacaacttg 450

gattctgata tccgtagagg tttggagagc aacgtgaatg ttgaattact 500

gaatgcttta catagtcaca tgattaataa gagaatgttg accaaggact 550

taaaaaatgg catgattatt ccttcaatgt ataacaattt ggggcttttc 600

attaaccatt atcctaatgg ggttgtcact gttaattgtg ctcgaatcat 650

ccatgggaac cagattgcaa caaatggtgt tgtccatgtc attgaccgtg 700

tgcttacaca aattggtacc tcaattcaag acttcattga agcagaagat 750

gacctttcat cttttagagc agctgccatc acatcggaca tattggaggc 800

ccttggaaga gacggtcact tcacactctt tgctcccacc aatgaggctt 850

ttgagaaact tccacgaggt gtcctagaaa ggatcatggg agacaaagtg 900

gcttccgaag ctcttatgaa gtaccacatc ttaaatactc tccagtgttc 950

tgagtctatt atgggaggag cagtctttga gacgctggaa ggaaatacaa 1000

ttgagatagg atgtgacggt gacagtataa cagtaaatgg aatcaaaatg 1050

gtgaacaaaa aggatattgt gacaaataat ggtgtgatcc atttgattga 1100

tcaggtccta attcctgatt ctgccaaaca agttattgag ctggctggaa 1150

aacagcaaac caccttcacg gatcttgtgg cccaattagg cttggcatct 1200

gctctgaggc cagatggaga atacactttg ctggcacctg tgaataatgc 1250

attttctgat gatactctca gcatggatca gcgcctcctt aaattaattc 1300

tgcagaatca catattgaaa gtaaaagttg gccttaatga gctttacaac 1350

gggcaaatac tggaaaccat cggaggcaaa cagctcagag tcttcgtata 1400

tcgtacagct gtctgcattg aaaattcatg catggagaaa gggagtaagc 1450

aagggagaaa cggtgcgatt cacatattcc gcgagatcat caagccagca 1500

gagaaatccc tccatgaaaa gttaaaacaa gataagcgct ttagcacctt 1550

cctcagccta cttgaagctg cagacttgaa agagctcctg acacaacctg 1600

gagactggac attatttgtg ccaaccaatg atgcttttaa gggaatgact 1650

agtgaagaaa aagaaattct gatacgggac aaaaatgctc ttcaaaacat 1700

cattctttat cacctgacac caggagtttt cattggaaaa ggatttgaac 1750

ctggtgttac taacatttta aagaccacac aaggaagcaa aatctttctg 1800

aaagaagtaa atgatacact tctggtgaat gaattgaaat caaaagaatc 1850

tgacatcatg acaacaaatg gtgtaattca tgttgtagat aaactcctct 1900

atccagcaga cacacctgtt ggaaatgatc aactgctgga aatacttaat 1950

aaattaatca aatacatcca aattaagttt gttcgtggta gcaccttcaa 2000

agaaatcccc gtgactgtct ataagccaat tattaaaaaa tacaccaaaa 2050

tcattgatgg agtgcctgtg gaaataactg aaaaagagac acgagaagaa 2100

cgaatcatta caggtcctga aataaaatac actaggattt ctactggagg 2150

tggagaaaca gaagaaactc tgaagaaatt gttacaagaa gacacacccg 2200

tgaggaagtt gcaagccaac aaaaaagttc aaggatctag aagacgatta 2250

agggaaggtc gttctcagtg aagggc 2276

<210> 44
<211> 2366
<212> DNA
<213> Homo sapiens

<400> 44
cccttgagac tcaagatgat tccctttta cccatgtttt ctctactatt 50

gctgcttatt gttaacccta taaacgccaa caatcattat gacaagatct 100

tggctcatag tcgtatcagg ggtcgggacc aaggcccaaa tgtctgtgcc 150

cttcaacaga ttttgggcac caaaaagaaa tacttcagca cttgtaagaa 200

ctggtataaa aagtccatct gtggacagaa aacgactgtg ttatatgaat 250

gttgccctgg ttatatgaga atggaaggaa tgaaaggctg cccagcagtt 300

ttgcccattg accatgttta tggcactctg ggcatcgtgg gagccaccac 350

aacgcagcgc tattctgacg cctcaaaact gagggaggag atcgagggaa 400

agggatcctt cacttacttt gcaccgagta atgaggcttg ggacaacttg 450

gattctgata tccgtagagg tttggagagc aacgtgaatg ttgaattact 500

gaatgcttta catagtcaca tgattaataa gagaatgttg accaaggact 550

taaaaaatgg catgattatt ccttcaatgt ataacaattt ggggcttttc 600

attaaccatt atcctaatgg ggttgtcact gttaattgtg ctcgaatcat 650

ccatgggaac cagattgcaa caaatggtgt tgtccatgtc attgaccgtg 700

tgcttacaca aattggtacc tcaattcaag acttcattga agcagaagat 750

gacctttcat cttttagagc agctgccatc acatcggaca tattggaggc 800

ccttggaaga gacggtcact tcacactctt tgctcccacc aatgaggctt 850

ttgagaaact tccacgaggt gtcctagaaa ggatcatggg agacaaagtg 900

gcttccgaag ctcttatgaa gtaccacatc ttaaatactc tccagtgttc 950

tgagtctatt atgggaggag cagtctttga gacgctggaa ggaaatacaa 1000

ttgagatagg atgtgacggt gacagtataa cagtaaatgg aatcaaaatg 1050

gtgaacaaaa aggatattgt gacaaataat ggtgtgatcc atttgattga 1100

tcaggtccta attcctgatt ctgccaaaca agttattgag ctggctggaa 1150

aacagcaaac caccttcacg gatcttgtgg cccaattagg cttggcatct 1200

gctctgaggc cagatggaga atacactttg ctggcacctg tgaataatgc 1250

attttctgat gatactctca gcatggatca gcgcctcctt aaattaattc 1300

tgcagaatca catattgaaa gtaaaagttg gccttaatga gctttacaac 1350

gggcaaatac tggaaaccat cggaggcaaa cagctcagag tcttcgtata 1400

tcgtacagct gtctgcattg aaaattcatg catggagaaa gggagtaagc 1450

aagggagaaa cggtgcgatt cacatattcc gcgagatcat caagccagca 1500

gagaaatccc tccatgaaaa gttaaaacaa gataagcgct ttagcacctt 1550

cctcagccta cttgaagctg cagacttgaa agagctcctg acacaacctg 1600

gagactggac attatttgtg ccaaccaatg atgcttttaa gggaatgact 1650

agtgaagaaa aagaaattct gatacgggac aaaaatgctc ttcaaaacat 1700

cattctttat cacctgacac caggagtttt cattggaaaa ggatttgaac 1750

ctggtgttac taacatttta aagaccacac aaggaagcaa aatctttctg 1800

aaagaagtaa atgatacact tctggtgaat gaattgaaat caaaagaatc 1850

tgacatcatg acaacaaatg gtgtaattca tgttgtagat aaactcctct 1900

atccagcaga cacacctgtt ggaaatgatc aactgctgga aatacttaat 1950

aaattaatca aatacatcca aattaagttt gttcgtggta gcaccttcaa 2000

agaaatcccc gtgactgtct atagacccac actaacaaaa gtcaaaattg 2050

aaggtgaacc tgaattcaga ctgattaaag aaggtgaaac aataactgaa 2100

gtgatccatg gagagccaat tattaaaaaa tacaccaaaa tcattgatgg 2150

agtgcctgtg gaaataactg aaaaagagac acgagaagaa cgaatcatta 2200

caggtcctga aataaaatac actaggattt ctactggagg tggagaaaca 2250

gaagaaactc tgaagaaatt gttacaagaa gacacacccg tgaggaagtt 2300

gcaagccaac aaaaaagttc aaggatctag aagacgatta agggaaggtc 2350

gttctcagtg aagggc 2366

<210> 45
<211> 2360
<212> DNA
<213> Homo sapiens

<400> 45
cccttgagac tcaagatgat tccctttta cccatgtttt ctctactatt 50

gctgcttatt gttaacccta taaacgccaa caatcattat gacaagatct 100

tggctcatag tcgtatcagg ggtcgggacc aaggcccaaa tgtctgtgcc 150

cttcaacaga ttttgggcac caaaaagaaa tacttcagca cttgtaagaa 200

ctggtataaa aagtccatct gtggacagaa aacgactgtg ttatatgaat 250

gttgccctgg ttatatgaga atggaaggaa tgaaaggctg cccagcagtt 300

ttgcccattg accatgttta tggcactctg ggcatcgtgg gagccaccac 350

aacgcagcgc tattctgacg cctcaaaact gagggaggag atcgagggaa 400

agggatcctt cacttacttt gcaccgagta atgaggcttg ggacaacttg 450

gattctgata tccgtagagg tttggagagc aacgtgaatg ttgaattact 500

gaatgcttta catagtcaca tgattaataa gagaatgttg accaaggact 550

taaaaaatgg catgattatt ccttcaatgt ataacaattt ggggcttttc 600

attaaccatt atcctaatgg ggttgtcact gttaattgtg ctcgaatcat 650

ccatgggaac cagattgcaa caaatggtgt tgtccatgtc attgaccgtg 700

tgcttacaca aattggtacc tcaattcaag acttcattga agcagaagat 750

gacctttcat cttttagagc agctgccatc acatcggaca tattggaggc 800

ccttggaaga gacggtcact tcacactctt tgctcccacc aatgaggctt 850

ttgagaaact tccacgaggt gtcctagaaa ggatcatggg agacaaagtg 900

gcttccgaag ctcttatgaa gtaccacatc ttaaatactc tccagtgttc 950

tgagtctatt atgggaggag cagtctttga gacgctggaa ggaaatacaa 1000

ttgagatagg atgtgacggt gacagtataa cagtaaatgg aatcaaaatg 1050

gtgaacaaaa aggatattgt gacaaataat ggtgtgatcc atttgattga 1100

tcaggtccta attcctgatt ctgccaaaca agttattgag ctggctggaa 1150

aacagcaaac caccttcacg gatcttgtgg cccaattagg cttggcatct 1200

gctctgaggc cagatggaga atacactttg ctggcacctg tgaataatgc 1250

attttctgat gatactctca gcatggatca gcgcctcctt aaattaattc 1300

tgcagaatca catattgaaa gtaaaagttg gccttaatga gctttacaac 1350

gggcaaatac tggaaaccat cggaggcaaa cagctcagag tcttcgtata 1400

tcgtacagct gtctgcattg aaaattcatg catggagaaa gggagtaagc 1450

aagggagaaa cggtgcgatt cacatattcc gcgagatcat caagccagca 1500

gagaaatccc tccatgaaaa gttaaaacaa gataagcgct ttagcacctt 1550

cctcagccta cttgaagctg cagacttgaa agagctcctg acacaacctg 1600

gagactggac attatttgtg ccaaccaatg atgcttttaa gggaatgact 1650

agtgaagaaa aagaaattct gatacgggac aaaaatgctc ttcaaaacat 1700

cattctttat cacctgacac caggagtttt cattggaaaa ggatttgaac 1750

ctggtgttac taacatttta aagaccacac aaggaagcaa aatctttctg 1800

aaagaagtaa atgatacact tctggtgaat gaattgaaat caaaagaatc 1850

tgacatcatg acaacaaatg gtgtaattca tgttgtagat aaactcctct 1900

atccagcaga cacacctgtt ggaaatgatc aactgctgga aatacttaat 1950

aaattaatca aatacatcca aattaagttt gttcgtggta gcaccttcaa 2000

agaaatcccc gtgactgtct ataagccaat tattaaaaaa tacaccaaaa 2050

tcattgatgg agtgcctgtg gaaataactg aaaaagagac acgagaagaa 2100

cgaatcatta caggtcctga aataaaatac actaggattt ctactggagg 2150

tggagaaaca gaagaaactc tgaagaaatt gttacaagaa gaggtcacca 2200

aggtcaccaa attcattgaa ggtggtgatg gtcatttatt tgaagatgaa 2250

gaaattaaaa gactgcttca gggagacaca cccgtgagga agttgcaagc 2300

caacaaaaaa gttcaaggat ctagaagacg attaagggaa ggtcgttctc 2350

agtgaagggc 2360

<210> 46
<211> 2186
<212> DNA
<213> Homo sapiens

<400> 46
acacggacca aggagtctaa cacgtgcgcg agtcgggggc tcgcacgaaa 50

gccgccgtgg cgcaatgaag gtgaaggccg gcgcgctcgc cggccgaggt 100

gggatcccga ggcctctcca gtccgccgag ggcgcaccac cggcccgtct 150

cgcccgccgc gccggggagg tggagcacga gcgcacgtgt taggacccga 200

aagatggtga actatgcctg ggcagggcga agccagagga aactctggtg 250

gaggtccgta gcggtcctga cgtgcaaatc ggtcgtccga cctgggtata 300

ggggcgggct ccaggcgagg cggtcgacgc tcctgaaaac ttgcgcgcgc 350

gctcgcgcca ctgcgcccgg agcgatgaag atggtcgcgc cctggacgcg 400

gttctactcc aacagctgct gcttgtgctg ccatgtccgc accggcacca 450

tcctgctcgg cgtctggtat ctgatcatca atgctgtggt actgttgatt 500

ttattgagtg ccctggctga tccggatcag tataactttt caagttctga 550

actgggaggt gactttgagt tcatggatga tgccaacatg tgcattgcca 600

ttgcgatttc tcttctcatg atcctgatat gtgctatggc tacttacgga 650

gcgtacaagc aacgcgcagc ctggatcatc ccattcttct gttaccagat 700

ctttgacttt gccctgaaca tgttggttgc aatcactgtg cttatttatc 750

caaactccat tcaggaatac atacggcaac tgcctcctaa ttttccctac 800

agagatgatg tcatgtcagt gaatcctacc tgtttggtcc ttattattct 850

tctgtttatt agcattatct tgacttttaa gggttacttg attagctgtg 900

tttggaactg ctaccgatac atcaatggta ggaactcctc tgatgtcctg 950

gtttatgtta ccagcaatga cactacggtg ctgctacccc cgtatgatga 1000

tgccactgtg aatggtgctg ccaaggagcc accgccacct tacgtgtctg 1050

cctaagcctt caagtgggcg gagctgaggg cagcagcttg actttgcaga 1100

catctgagca atagttctgt tatttcactt ttgccatgag cctctctgag 1150

cttgtttgtt gctgaaatgc tacttttaa aatttagatg ttagattgaa 1200

aactgtagtt ttcaacatat gctttgctag aacactgtga tagattaact 1250

gtagaattct tcctgtacga ttggggatat aatgggcttc actaaccttc 1300

cctaggcatt gaaacttccc ccaaatctga tggacctaga agtctgcttt 1350

tgtacctgct gggccccaaa gttgggcatt tttctctctg ttccctctct 1400

tttgaaaatg taaaataaaa ccaaaaatag acaacttttt cttcagccat 1450

tccagcatag agaacaaaac cttatggaaa caggaatgtc aattgtgtaa 1500

tcattgttct aattaggtaa atagaagtcc ttatgtatgt gttacaagaa 1550

tttcccccac aacatccttt atgactgaag ttcaatgaca gtttgtgttt 1600

gggtggtaaa ggattttctc catggcctga attaagacca ttagaaagca 1650

ccaggccgtg ggagcagtga ccatctgctg actgttcttg tggatcttgt 1700

gtccagggac atggggtgac atgcctcgta tgtgttagag ggtggaatgg 1750

atgtgtttgg cgctgcatgg gatctggtgc ccctcttctc ctggattcac 1800

atccccaccc agggcccgct tttactaagt gttctgccct agattggttc 1850

aaggaggtca tccaactgac tttatcaagt ggaattggga tatatttgat 1900

atacttctgc ctaacaacat ggaaaagggt tttcttttcc ctgcaagcta 1950

catcctactg ctttgaactt ccaagtatgt ctagtcacct tttaaaatgt 2000

aaacattttc agaaaaatga ggattgcctt ccttgtatgc gctttttacc 2050

ttgactacct gaattgcaag ggatttttat atattcatat gttacaaagt 2100

cagcaactct cctgttggtt cattattgaa tgtgctgtaa attaagttgt 2150

ttgcaattaa aacaaggttt gcccacaaaa aaaaaa 2186

<210> 47
<211> 851
<212> DNA
<213> Homo sapiens

<400> 47
attcgccctt ccgaaagatg gtgaactatg cctgggcagg gcgggctcca 50

ggcgaggcgg tcgacgctcc tgaaaacttg cgcgcgcgct cgcgccactg 100

cgcccggagc gatgaagatg gtcgcgccct ggactcggtt ctactccaac 150

agctgctgct tgtgctgcca tgtccgcacc ggcaccatcc tgctcggcgt 200

ctggtatctg atcatcaatg ctgtggtact gttgatttta ttgagtgccc 250

tggctgatcc ggatcagtat aacttttcaa gttctgaact gggaggtgac 300

tttgagttca tggatgatgc caacatgtgc attgccattg cgatttctct 350

tctcatgatc ctgatatgtg ctatggctac ttacggagcg tacaagcaac 400

gcgcagcctg gatcatccca ttcttctgtt accagatctt tgactttgcc 450

ctgaacatgt tggttgcaat cactgtgctt atttatccaa actccattca 500

ggaatacata cggcaactgc ctcctaattt tccctacaga gatgatgtca 550

tgtcagtgaa tcctacctgt ttggtcctta ttattcttct gtttattagc 600

attatcttga cttttaaggg ttacttgatt agctgtgttt ggaactgcta 650

ccgatacatc aatggtagga actcctctga tgtcctggtt tatgttacca 700

gcaatgacac tacggtgctg ctacccccgt atgatgatgc cactgtgaat 750

ggtgctgcca aggagccacc gccaccttac gtgtctgcct aagccttcaa 800

gtgggcggag ctgagggcag cagcttgact ttgcagacat ctgagaaggg 850

c 851

<210> 48
<211> 1728
<212> DNA
<213> Homo sapiens

<400> 48
agcgccatgg ggcgagccgg gggcgggggc ccgggccggg ggccgccgcc 50

actgctgctg tttctggggg ccgcgctggt cctggcctct ggggccgtgc 100

cggcgcgtga ggcgggcagc gcggttgagg ccgaagagct ggtgaagggc 150

agcccggcgt gggagccgcc tgccaacgac acgcgggaag aagccggccc 200

accagcggct ggggaagatg aggcgtcgtg gacggcgccc ggtggcgagc 250

tggccgggcc agaagaggtg ctgcaggagt cggctgcggt gaccggcacc 300

gcctggctgg aagctgacag cccaggcctg ggaggagtga ccgcagaggc 350

gggcagcggc gatgcccagg cccttccagc tacgctccag gctccccacg 400

aggtcctcgg gcagtcaatc atgccccctg ccattcctga ggctacagag 450

gccagcgggc caccctcccc caccccggc gacaagctga gcccagcttc 500

tgaactcccc aaggagagcc ccttggaggt ttggctgaac ctgggggggca 550

gcacacccga ccctcaagtg ccagagctga cttacccatt tcagggcacc 600

ctggagcccc aaccggcatc agatatcatt gacatcgact acttcgaagg 650

actggatggt gagggtcgtg gcgcagatct ggggagcttc ccagggtcac 700

caggaacctc agagaaccac cctgatactg agggagagac cccttcctgg 750

agcctgcttg acttatacga tgatttcacc cccttcgatg aatctgattt 800

ctaccccacc acatcctttt atgatgactt ggatgaagag gaggaggaag 850

aggaggatga caaagatgca gtaggaggtg gagacctaga agatgaaaat 900

gagcttctag tgcccactgg gaagcctggt ctggggcccg ggacaggcca 950

gcccaccagt cggtggcatg ctgtccctcc acagcacact ctggggtcgg 1000

tccccggcag cagcatcgcc ctcaggcccc gcccaggaga gccaggcagg 1050

gacttggcct ccagtgaaaa tggcactgag tgccgcagtg gctttgtgcg 1100

gcataacggc tcctgccggt cagtgtgcga cctcttccca agttactgtc 1150

acaatggcgg ccagtgctac ctggtggaga acataggggc cttctgcagg 1200

tgcaacacgc aggactacat ctggcacaag gggatgcgct gcgagtccat 1250

catcaccgac ttccaggtga tgtgcgtggc cgtgggctcg gctgccctcg 1300

tcctgctcct gctcttcatg atgacggtgt tctttgccaa gaagctctac 1350

ctgctcaaga cggagaatac caagctgcgt aggaccaaca aattccggac 1400

cccatctgag ctccacaatg ataacttctc cctctccacc attgccgagg 1450

gctctcaccc aaatgatgat cctagtgctc cccacaaaat ccaggaggtt 1500

ctcaagtcct gcctgaaaga ggaggagtca tttaacatcc agaactccat 1550

gtcgcccaaa cttgagggtg gcaaaggtga ccaggctgac ttggatgtga 1600

actgtcttca gaataattta acctaaagca gagcaagaag agaggaagcg 1650

gggtagtggg tgggggtagg gaagaaacat tatctcctct tgtacagagt 1700

ctatttcttg taaccatttg ttaaactc 1728

<210> 49
<211> 4231
<212> DNA
<213> Homo sapiens

<400> 49
ggacaggcgt ggcggccgga gccccagcat ccctgcttga ggtccaggag 50

cggagcccgc ggccaccgcc gcctgatcag cgcgaccccg gcccgcgccc 100

gccccgcccg gcaagatgct gcccgtgtac caggaggtga agcccaaccc 150

gctgcaggac gcgaacatct gctcacgcgt gttcttctgg tggctcaatc 200

ccttgtttaa aattggccat aaacggagat tagaggaaga tgatatgtat 250

tcagtgctgc cagaagaccg ctcacagcac cttggagagg agttgcaagg 300

gttctgggat aaagaagttt taagagctga gaatgacgca cagaagcctt 350

ctttaacaag agcaatcata aagtgttact ggaaatctta tttagttttg 400

ggaattttta cgttaattga ggaaagtgcc aaagtaatcc agcccatatt 450

tttgggaaaa attattaatt attttgaaaa ttatgatccc atggattctg 500

tggctttgaa cacagcgtac gcctatgcca cggtgctgac tttttgcacg 550

ctcattttgg ctatactgca tcacttatat ttttatcacg ttcagtgtgc 600

tgggatgagg ttacgagtag ccatgtgcca tatgatttat cggaaggcac 650

ttcgtcttag taacatggcc atggggaaga caaccacagg ccagatagtc 700

aatctgctgt ccaatgatgt gaacaagttt gatcaggtga cagtgttctt 750

acacttcctg tgggcaggac cactgcaggc gatcgcagtg actgccctac 800

tctggatgga gataggaata tcgtgccttg ctgggatggc agttctaatc 850

attctcctgc ccttgcaaag ctgttttggg aagttgttct catcactgag 900

gagtaaaact gcaactttca cggatgccag gatcaggacc atgaatgaag 950

ttataactgg tataaggata ataaaaatgt acgcctggga aaagtcattt 1000

tcaaatctta ttaccaattt gagaaagaag gagatttcca agattctgag 1050

aagttcctgc ctcaggggga tgaatttggc ttcgtttttc agtgcaagca 1100

aaatcatcgt gtttgtgacc ttcaccacct acgtgctcct cggcagtgtg 1150

atcacagcca gccgcgtgtt cgtggcagtg acgctgtatg gggctgtgcg 1200

gctgacggtt accctcttct tcccctcagc cattgagagg gtgtcagagg 1250

caatcgtcag catccgaaga atccagacct ttttgctact tgatgagata 1300

tcacagcgca accgtcagct gccgtcagat ggtaaaaaga tggtgcatgt 1350

gcaggatttt actgcttttt gggataaggc atcagagacc ccaactctac 1400

aaggcctttc ctttactgtc agacctggcg aattgttagc tgtggtcggc 1450

cccgtgggag cagggaagtc atcactgtta agtgccgtgc tcggggaatt 1500

ggccccaagt cacgggctgg tcagcgtgca tggaagaatt gcctatgtgt 1550

ctcagcagcc ctgggtgttc tcgggaactc tgaggagtaa tattttattt 1600

gggaagaaat atgaaaagga acgatatgaa aaagtcataa aggcttgtgc 1650

tctgaaaaag gatttacagc tgttggagga tggtgatctg actgtgatag 1700

gagatcgggg aaccacgctg agtggagggc agaaagcacg ggtaaacctt 1750

gcaagagcag tgtatcaaga tgctgacatc tatctcctgg acgatcctct 1800

cagtgcagta gatgcggaag ttagcagaca cttgttcgaa ctgtgtattt 1850

gtcaaatttt gcatgagaag atcacaattt tagtgactca tcagttgcag 1900

tacctcaaag ctgcaagtca gattctgata ttgaaagatg gtaaaatggt 1950

gcagaagggg acttacactg agttcctaaa atctggtata gattttggct 2000

ccctttaaa gaaggataat gaggaaagtg aacaacctcc agttccagga 2050

actcccacac taaggaatcg taccttctca gagtcttcgg tttggtctca 2100

acaatcttct agaccctcct tgaaagatgg tgctctggag agccaagata 2150

cagagaatgt cccagttaca ctatcagagg agaaccgttc tgaaggaaaa 2200

gttggttttc aggcctataa gaattacttc agagctggtg ctcactggat 2250

tgtcttcatt ttccttattc tcctaaacac tgcagctcag gttgcctatg 2300

tgcttcaaga ttggtggctt tcatactggg caaacaaaca aagtatgcta 2350

aatgtcactg taaatggagg aggaaatgta accgagaagc tagatcttaa 2400

ctggtactta ggaatttatt caggtttaac tgtagctacc gttctttttg 2450

gcatagcaag atctctattg gtattctacg tccttgttaa ctcttcacaa 2500

actttgcaca acaaaatgtt tgagtcaatt ctgaaagctc cggtattatt 2550

ctttgataga aatccaatag gaagaatttt aaatcgtttc tccaaagaca 2600

ttggacactt ggatgatttg ctgccgctga cgtttttaga tttcatccag 2650

acattgctac aagtggttgg tgtggtctct gtggctgtgg ccgtgattcc 2700

ttggatcgca ataccccttgg ttccccttgg aatcattttc atttttcttc 2750

ggcgatattt tttggaaacg tcaagagatg tgaagcgcct ggaatctaca 2800

actcggagtc cagtgttttc ccacttgtca tcttctctcc aggggctctg 2850

gaccatccgg gcatacaaag cagaagagag gtgtcaggaa ctgtttgatg 2900

cacaccagga tttacattca gaggcttggt tcttgttttt gacaacgtcc 2950

cgctggttcg ccgtccgtct ggatgccatc tgtgccatgt ttgtcatcat 3000

cgttgccttt gggtccctga ttctggcaaa aactctggat gccgggcagg 3050

ttggtttggc actgtcctat gccctcacgc tcatggggat gtttcagtgg 3100

tgtgttcgac aaagtgctga agttgagaat atgatgatct cagtagaaag 3150

ggtcattgaa tacacagacc ttgaaaaaga agcaccttgg gaatatcaga 3200

aacgcccacc accagcctgg ccccatgaag gagtgataat ctttgacaat 3250

gtgaacttca tgtacagtcc aggtgggcct ctggtactga agcatctgac 3300

agcactcatt aaatcacaag aaaaggttgg cattgtggga agaaccggag 3350

ctggaaaaag ttccctcatc tcagcccttt ttagattgtc agaacccgaa 3400

ggtaaaattt ggattgataa gatcttgaca actgaaattg gacttcacga 3450

tttaaggaag aaaatgtcaa tcatacctca ggaacctgtt ttgttcactg 3500

gaacaatgag gaaaaacctg gatcccttta aggagcacac ggatgaggaa 3550

ctgtggaatg ccttacaaga ggtacaactt aaagaaacca ttgaagatct 3600

tcctggtaaa atggatactg aattagcaga atcaggatcc aattttagtg 3650

ttggacaaag acaactggtg tgccttgcca gggcaattct caggaaaaat 3700

cagatattga ttattgatga agcgacggca aatgtggatc caagaactga 3750

tgagttaata caaaaaaaaa tccgggagaa atttgcccac tgcaccgtgc 3800

taaccattgc acacagattg aacaccatta ttgacagcga caagataatg 3850

gttttagatt caggaagact gaaagaatat gatgagccgt atgttttgct 3900

gcaaaataaa gagagcctat tttacaagat ggtgcaacaa ctgggcaagg 3950

cagaagccgc tgccctcact gaaacagcaa aacaggtata cttcaaaaga 4000

aattatccac atattggtca cactgaccac atggttacaa acacttccaa 4050

tggacagccc tcgaccttaa ctattttcga gacagcactg tgaatccaac 4100

caaaatgtca agtccgttcc gaaggcattt tccactagtt tttggactat 4150

gtaaaccaca ttgtactttt ttttactttg gcaacaaata tttatacata 4200

caagatgcta gttcatttga atatttctcc c 4231

<210> 50
<211> 2861
<212> DNA
<213> Homo sapiens

<400> 50
ctccgtccca ggggagaagg agaggcgtct gcaggggca gagaccgcag 50

ctacctgccg ggtgcgcccc ccacccagga gcgctcgctt cgccccttt 100

cctccccccg cccccacctc cttattggtg ctagtttgca gcgcccagct 150

cctgcgcctt cgcttcgcgt ttgaatctgg ctcgcccctt cgtattatgt 200

ctgcactccg aaggaaattt ggggacgatt atcaggtagt gaccacatcg 250

tccagcggct cgggcttgca gccccagggg ccaggccagg accctcagca 300

gcagcttgtg cccaagaaga agcggcagcg gttcgtggac aagaacggcc 350

ggtgcaatgt acagcacggc aacctgggca gcgagacaag ccgctacctc 400

tcggacctct tcaccacgct ggtggacctc aagtggcgct ggaacctctt 450

catcttcatt ctcacctaca ccgtggcctg gcttttcatg gcgtccatgt 500

ggtgggtgat cgcctacact cggggcgacc tgaacaaagc ccacgtcggt 550

aactacacgc cttgcgtggc caatgtctat aacttccctt ctgccttcct 600

cttcttcatc gagacggagg ccaccatcgg ctatggctac cgatacatca 650

cagacaagtg ccccgagggc atcatcctct tcctcttcca gtccatcctg 700

ggctccatcg tggacgcctt cctcatcggc tgcatgttca tcaagatgtc 750

ccagcccaag aagcgcgccg agaccctcat gttcagcgag cacgcggtga 800

tctccatgag ggacggaaaa ctcacgctta tgttccgggt gggcaacctg 850

cgcaacagcc acatggtctc cgcgcagatt cgctgcaagc tgctcaaatc 900

tcggcagaca cctgagggtg agttccttcc ccttgaccaa cttgaactgg 950

atgtaggttt tagtacaggg gcagatcaac tttttcttgt gtccccctc 1000

acaatttgcc acgtgatcga tgccaaaagc ccctttatg acctatccca 1050

gcgaagcatg caaactgaac agttcgagat tgtcgtcatc ctagaaggca 1100

ttgtggaaac aactgggatg acttgtcaag ctcgaacatc atatactgaa 1150

gatgaagttc tttggggtca tcgttttttt cctgtaattt ccttagaaga 1200

gggattcttt aaagttgatt actcccagtt ccacgcaaca tttgaagtcc 1250

ccaccccacc ttacagtgtg aaagagcagg aggaaatgct tctcatgtcg 1300

tccccttaa tagcaccagc cataactaac agcaaagaaa gacataattc 1350

tgtggaatgc ttagatggac tagatgatat tactacaaaa ctaccatcta 1400

agctgcagaa aattactgga agagaagact ttcccaaaaa actcttgagg 1450

atgagttcta caacttcaga aaaagcctac agcttgggag acttgcccat 1500

gaaacttcaa cgaataagtt cagttccggg caactcagaa gaaaaactgg 1550

tatctaaaac caccaagatg ttatctgatc ccatgagcca gtctgtggct 1600

gatttgccac caaagcttca aaagatggct ggaggagcag ctaggatgga 1650

agggaacctt ccagccaaat taagaaaaat gaactctgat cgcttcacat 1700

aacaaagcac tcccttaggc attatttaat gtttgattta gtaatagtcc 1750

aatatttggc gatgaggtaa ttctccctaa ggaatctgaa agtatatttt 1800

cctcccagtt ctacagcata tttcgagaac ccttcctttc ccaagtattg 1850

cgaatgtgca gaaagcaaca gttacggagg gaggacatca taaggaagtt 1900

attaacgggc atgtattatc acatcaagca tgcaataatg tgcaaatttt 1950

gcatttagtt ttatggcatg atttatatat ggcatattta tattgtatat 2000

tctggaaaaa aaaatatata tatatattta aaggggagat actctccctg 2050

acatttctaa catatgtatt aagccaaaca tgagtgaata gctttcaggg 2100

cgataaaact aaatatatgt ctgtgtgtgt gtgtgtatgt atacacacat 2150

atacatatat atatacacat acatacacat acatacatac atacatatat 2200

atctgataaa attgtgatgt tttgttcaaa gttgtagttc ttgtgcatgt 2250

ttactttatt agagtaggaa ggctactggc attaattatt aataccaaat 2300

attttagcct taaatttttg tcattttaaa atctgattta atgttttctg 2350

ctgtttaagg tcttgggagg ctttcaattg tattttatat gagagaatca 2400

cacaagtttg tgctatctat ggccctgcaa aaatataacc attacatgtt 2450

taaattgtaa attttagagc ataccagtac tcagtatagc attgaacatt 2500

tcttatgatt tttaaaagtt gctagtactg gggagaaata attgttgatt 2550

aatttgagaa ttattccttt cctagactaa ttaaaatctg gaaatctgtt 2600

ttgtatatga tctaatacaa agatgagctc tgaacaaaca ctgaatcatg 2650

ttaatagaca gtagccaagt tatattgaat atatcagaat ctgtgtgaag 2700

ttacacaatt aattgtccct gtttcaaact gagtaaattg gaaacatttt 2750

ctttcttttt ctggaaattt tgtccatttt aaaaaccaat cattttaaga 2800

agacatgaca atgcaatgaa acagatgata aatatttatg cttaaaataa 2850

aaaaaaaaaa a 2861

<210> 51
<211> 3557
<212> DNA
<213> Homo sapiens

<400> 51
gagagggtcc ttcagggtct gcttatgccc ttgttcaaga acaccagtgt 50

cagctctctg tactctggtt gcagactgac cttgctcagg cctgagaagg 100

atggggcagc caccagagtg gatgctgtct gcacccatcg tcctgacccc 150

aaaagccctg gactggacag agagcggctg tactggaagc tgagccagct 200

gacccacggc atcactgagc tgggccccta caccctggac aggcacagtc 250

tctatgtcaa tggtttcacc catcagagct ctatgacgac caccagaact 300

cctgatacct ccacaatgca cctggcaacc tcgagaactc cagcctccct 350

gtctggacct acgaccgcca gccctctcct ggtgctattc acaattaact 400

tcaccatcac taacctgcgg tatgaggaga acatgcatca ccctggctct 450

agaaagttta acaccacgga gagagtcctt cagggtctgc tcaggcctgt 500

gttcaagaac accagtgttg gccctctgta ctctggctgc agactgacct 550

tgctcaggcc caagaaggat ggggcagcca ccaaagtgga tgccatctgc 600

acctaccgcc ctgatcccaa aagccctgga ctggacagag agcagctata 650

ctgggagctg agccagctaa cccacagcat cactgagctg ggcccctaca 700

ccctggacag ggacagtctc tatgtcaatg gtttcacaca gcggagctct 750

gtgcccacca ctagcattcc tgggaccccc acagtggacc tgggaacatc 800

tgggactcca gtttctaaac ctggtccctc ggctgccagc cctctcctgg 850

tgctattcac tctcaacttc accatcacca acctgcggta tgaggagaac 900

atgcagcacc ctggctccag gaagttcaac accacggaga gggtccttca 950

gggcctgctc aggtccctgt tcaagagcac cagtgttggc cctctgtact 1000

ctggctgcag actgactttg ctcaggcctg aaaaggatgg gacagccact 1050

ggagtggatg ccatctgcac ccaccaccct gaccccaaaa gccctaggct 1100

ggacagagag cagctgtatt gggagctgag ccagctgacc cacaatatca 1150

ctgagctggg ccactatgcc ctggacaacg acagcctctt tgtcaatggt 1200

ttcactcatc ggagctctgt gtccaccacc agcactcctg ggacccccac 1250

agtgtatctg ggagcatcta agactccagc ctcgatattt ggcccttcag 1300

ctgccagcca tctcctgata ctattcaccc tcaacttcac catcactaac 1350

ctgcggtatg aggagaacat gtggcctggc tccaggaagt tcaacactac 1400

agagagggtc cttcagggcc tgctaaggcc cttgttcaag aacaccagtg 1450

ttggccctct gtactctggc tccaggctga ccttgctcag gccagagaaa 1500

gatggggaag ccaccggagt ggatgccatc tgcacccacc gccctgaccc 1550

cacaggccct gggctggaca gagagcagct gtatttggag ctgagccagc 1600

tgacccacag catcactgag ctgggcccct acacactgga cagggacagt 1650

ctctatgtca atggtttcac ccatcggagc tctgtaccca ccaccagcac 1700

cggggtggtc agcgaggagc cattcacact gaacttcacc atcaacaacc 1750

tgcgctacat ggcggacatg ggccaacccg gctccctcaa gttcaacatc 1800

acagacaacg tcatgaagca cctgctcagt cctttgttcc agaggagcag 1850

cctgggtgca cggtacacag gctgcagggt catcgcacta aggtctgtga 1900

agaacggtgc tgagacacgg gtggacctcc tctgcaccta cctgcagccc 1950

ctcagcggcc caggtctgcc tatcaagcag gtgttccatg agctgagcca 2000

gcagacccat ggcatcaccc ggctgggccc ctactctctg gacaaagaca 2050

gcctctacct taacggttac aatgaacctg gtctagatga gcctcctaca 2100

actcccaagc cagccaccac attcctgcct cctctgtcag aagccacaac 2150

agccatgggg taccacctga agaccctcac actcaacttc accatctcca 2200

atctccagta ttcaccagat atgggcaagg gctcagctac attcaactcc 2250

accgaggggg tccttcagca cctgctcaga cccttgttcc agaagagcag 2300

catgggcccc ttctacttgg gttgccaact gatctccctc aggcctgaga 2350

aggatggggc agccactggt gtggacacca cctgcaccta ccaccctgac 2400

cctgtgggcc ccgggctgga catacagcag ctttactggg agctgagtca 2450

gctgacccat ggtgtcaccc aactgggctt ctatgtcctg gacagggata 2500

gcctcttcat caatggctat gcaccccaga atttatcaat ccggggcgag 2550

taccagataa atttccacat tgtcaactgg aacctcagta atccagaccc 2600

cacatcctca gagtacatca ccctgctgag ggacatccag gacaaggtca 2650

ccacactcta caaaggcagt caactacatg acacattccg cttctgcctg 2700

gtcaccaact tgacgatgga ctccgtgttg gtcactgtca aggcattgtt 2750

ctcctccaat ttggacccca gcctggtgga gcaagtcttt ctagataaga 2800

ccctgaatgc ctcattccat tggctgggct ccacctacca gttggtggac 2850

atccatgtga cagaaatgga gtcatcagtt tatcaaccaa caagcagctc 2900

cagcacccag cacttctacc cgaatttcac catcaccaac ctaccatatt 2950

cccaggacaa agcccagcca ggcaccacca attaccagag gaacaaaagg 3000

aatattgagg atgcgctcaa ccaactcttc cgaaacagca gcatcaagag 3050

ttatttttct gactgtcaag tttcaacatt caggtctgtc cccaacaggc 3100

accacaccgg ggtggactcc ctgtgtaact tctcgccact ggctcggaga 3150

gtagacagag ttgccatcta tgaggaattt ctgcggatga cccggaatgg 3200

tacccagctg cagaacttca ccctggacag gagcagtgtc cttgtggatg 3250

ggtattctcc caacagaaat gagcccttaa ctgggaattc tgaccttccc 3300

ttctgggctg tcatcttcat cggcttggca ggactcctgg gactcatcac 3350

atgcctgatc tgcggtgtcc tggtgaccac ccgccggcgg aagaaggaag 3400

gagaatacaa cgtccagcaa cagtgcccag gctactacca gtcacaccta 3450

gacctggagg atctgcaatg actggaactt gccggtgcct ggggtgcctt 3500

tcccccagcc agggtccaaa gaagcttggc tggggcagaa ataaaccata 3550

ttggtcg 3557

<210> 52
<211> 3866
<212> DNA
<213> Homo sapiens

<400> 52
gcccgggcag gtgctgatag cacagttctg tccagagaag gaaggcggaa 50

taaacttatt cattcccagg aactcttggg gtaggtgtgt gtttttcaca 100

tcttaaaggc tcacagaccc tgcgctggac aaatgttcca ttcctgaagg 150

acctctccag aatccggatt gctgaatctt ccctgttgcc tagaagggct 200

ccaaaccacc tcttgacaat gggaaactgg gtggttaacc actggttttc 250

agttttgttt ctggttgttt ggttagggct gaatgttttc ctgtttgtgg 300

atgccttcct gaaatatgag aaggccgaca aatactacta cacaagaaaa 350

atccttgggt caacattggc ctgtgcccga gcgtctgctc tctgcttgaa 400

ttttaacagc acgctgatcc tgcttcctgt gtgtcgcaat ctgctgtcct 450

tcctgagggg cacctgctca ttttgcagcc gcacactgag aaagcaattg 500

gatcacaacc tcaccttcca caagctggtg gcctatatga tctgcctaca 550

tacagctatt cacatcattg cacacctgtt taactttgac tgctatagca 600

gaagccgaca ggccacagat ggctcccttg cctccattct ctccagccta 650

tctcatgatg agaaaaaggg gggttcttgg ctaaatccca tccagtcccg 700

aaacacgaca gtggagtatg tgacattcac cagcgttgct ggtctcactg 750

gagtgatcat gacaatagcc ttgattctca tggtaacttc agctactgag 800

ttcatccgga ggagttattt tgaagtcttc tggtatactc accacctttt 850

tatcttctat atccttggct tagggattca cggcattggt ggaattgtcc 900

ggggtcaaac agaggagagc atgaatgaga gtcatcctcg caagtgtgca 950

gagtcttttg agatgtggga tgatcgtgac tcccactgta ggcgccctaa 1000

gtttgaaggg catccccctg agtcttggaa gtggatcctt gcaccggtca 1050

ttctttatat ctgtgaaagg atcctccggt tttaccgctc ccagcagaag 1100

gttgtgatta ccaaggttgt tatgcaccca tccaaagttt tggaattgca 1150

gatgaacaag cgtggcttca gcatggaagt ggggcagtat atctttgtta 1200

attgcccctc aatctctctc ctggaatggc atcctttttac tttgacctct 1250

gctccagagg aagatttctt ctccattcat atccgagcag caggggactg 1300

gacagaaaat ctcataaggg ctttcgaaca acaatattca ccaattccca 1350

ggattgaagt ggatggtccc tttggcacag ccagtgagga tgttttccag 1400

tatgaagtgg ctgtgctggt tggagcagga attggggtca ccccctttgc 1450

ttctatcttg aaatccatct ggtacaaatt ccagtgtgca gaccacaacc 1500

tcaaaacaaa aaagatctat ttctactgga tctgcaggga gacaggtgcc 1550

ttttcctggt tcaacaacct gttgacttcc ctggaacagg agatggagga 1600

attaggcaaa gtgggttttc taaactaccg tctcttcctc accggatggg 1650

acagcaatat tgttggtcat gcagcattaa actttgacaa ggccactgac 1700

atcgtgacag gtctgaaaca gaaaacctcc tttgggagac caatgtggga 1750

caatgagttt tctacaatag ctacctccca ccccaagtct gtagtgggag 1800

ttttcttatg tggccctcgg actttggcaa agagcctgcg caaatgctgt 1850

caccgatatt ccagtctgga tcctagaaag gttcaattct acttcaacaa 1900

agaaaatttt tgagttatag gaataaggac ggtaatctgc attttgtctc 1950

tttgtatctt cagtaattga gttataggaa taaggacggt aatctgcatt 2000

ttgtctcttt gtatcttcag taatttactt ggtctcgtca ggtttgagca 2050

gtcactttag gataagaatg tgcctctcaa gccttgactc cctggtattc 2100

ttttttgat tgcattcaac ttcgttactt gagcttcagc aacttaagaa 2150

cttctgaagt tcttaaagtt ctgaagttct taaagcccat ggatcctttc 2200

tcagaaaaat aactgtaaat ctttctggac agccatgact gtagcaaggc 2250

ttgatagcag aggtttggtg gttcagagtt atacaactaa tcccaggtga 2300

ttttatcaat tccagtgtta ccatctcctg agttttggtt tgtaatcttt 2350

tgtccctccc acccccacag aagatttcta agtagggtga ctttttaaat 2400

aaaaatttat tgaataatta atgataaaac ataataataa acataaataa 2450

taaacaaaat taccgagaac cccatcccca tataacacca acagtgtaca 2500

tgtttactgt cacttttgat atggtcttat ccagtgtgaa cagcaattta 2550

ttatttttgc tcatcaaaaa ataaaggatt ttcttcttca cttgatgaat 2600

gagtcttttg ttacttcttt ttgggctctc catttgggaa agaaaatcta 2650

gaaagtatgc tcattgtaat tgaaattgta ttataattgc aagtcagcta 2700

tactgtcatc ttgtattatt ctgttaaaag cgattgtgaa aacttgaact 2750

ccatgttagg atgacttatt tagtcaagag gagctcttta tcactattaa 2800

gggagtttat taatgtctag tttaaaaaat aacaatttat aattttcata 2850

gatttttgtg gactcctacc tatatcgcta aggaatcatt cagctaataa 2900

cataagcaaa atgccttcat tctctatcag aatgcttttt atttctgcca 2950

tctacccaac ctcggcctac aaggaccaca ctggctctgg gatgttagtc 3000

gtcacagtga atccctagaa ggagagaagc ccagggctct taagagagaa 3050

acctcacact tcattgtgcc cagacaattc aactcattat agattagaag 3100

aatggttgta ctgaattagc attatattat ctatttttc taccattaag 3150

tgaccacgaa tgaacaaatg gttgtttcta atagtttagg aacctctacg 3200

gcttgaggga taaaactgtt tttggttttt ttttgtcact agtctgagca 3250

caataaaggt aggtcatgaa ctcatggtgg ctaatcatgt ttcatttatt 3300

acaacagtaa aacaatactg atggttgttt acatagggag ggataccact 3350

gctgttgcca atttgttcac ccggatttca gctaattcag gtttaataag 3400

gctttgggaa aatgcacaat gaaaccttga agacatgatt cttattagac 3450

aactactgac caaatgacaa cttcatcttc aatgaccttg atttggaatc 3500

ttgaaactta caacgagttc ctaatgagaa aggcatcact cttggtctcc 3550

atgtctaaat atattgacac actcacttca agatatgatg gctgccaaaa 3600

tcaatatccc aaggcttgaa tatttatggg agtaggcata aattcgtgat 3650

tatattcatg gtaatgaata taataaatcc tatgcatagc caacatccaa 3700

gttgaactgg ggatgagcag gatacacaga aagctcctca cctgctttga 3750

gcaagttttt cttgcaaagt tgtaatatga aataccttaa tcaacaaatt 3800

tcttataaat gtttatatga gtagaaaatc cctcaatttc attgtcactc 3850

aaaaagttga ttatct 3866

<210> 53
<211> 1817
<212> DNA
<213> Homo sapiens

<400> 53
cccttgggct ccaaaccacc tcttgacaat gggaaactgg gtggttaacc 50

actggttttc agttttgttt ctggttgttt ggttagggct gaatgttttc 100

ctgtttgtgg atgccttcct gaaatatgag aaggccgaca aatactacta 150

cacaagaaaa atccttgggt caacattggc ctgtgcccga gcgtctgctc 200

tctgcttgaa ttttaacagc acgctgatcc tgcttcctgt gtgtcgcaat 250

ctgctgtcct tcctgagggg cacctgctca ttttgcagcc gcacactgag 300

aaagcaattg gatcacaacc tcaccttcca caagctggtg gcctatatga 350

tctgcctaca tacagctatt cacatcattg cacacctgtt taactttgac 400

tgctatagca gaagccgaca ggccacagat ggctcccttg cctccattct 450

ctccagccta tctcatgatg agaaaaaggg gggttcttgg ctaaatccca 500

tccagtcccg aaacacgaca gtggagtatg tgacattcac cagcattgct 550

ggtctcactg gagtgatcat gacaatagcc ttgattctca tggtaacttc 600

agctactgag ttcatccgga ggagttattt tgaagtcttc tggtatactc 650

accacctttt tatcttctat atccttggct tagggattca cggcattggt 700

ggaattgtcc ggggtcaaac agaggagagc atgaatgaga gtcatcctcg 750

caagtgtgca gagtcttttg agatgtggga tgatcgtgac tcccactgta 800

ggcgccctaa gtttgaaggg catccccctg agtcttggaa gtggatcctt 850

gcaccggtca ttctttatat ctgtgaaagg atcctccggt tttaccgctc 900

ccagcagaag gttgtgatta ccaaggttgt tatgcaccca tccaaagttt 950

tggaattgca gatgaacaag cgtggcttca gcatggaagt ggggcagtat 1000

atctttgtta attgcccctc aatctctctc ctggaatggc atccttttac 1050

tttgacctct gctccagagg aagatttctt ctccattcat atccgagcag 1100

caggggactg gacagaaaat ctcataaggg ctttcgaaca acaatattca 1150

ccaattccca ggattgaagt ggatggtccc tttggcacag ccagtgagga 1200

tgttttccag tatgaagtgg ctgtgctggt tggagcagga attggggtca 1250

ccccctttgc ttctatcttg aaatccatct ggtacaaatt ccagtgtgca 1300

gaccacaacc tcaaaacaaa aaagatctat ttctactgga tctgcaggga 1350

gacaggtgcc ttttcctggt tcaacaacct gttgacttcc ctggaacagg 1400

agatggagga attaggcaaa gtgggttttc taaactaccg tctcttcctc 1450

accggatggg acagcaatat tgttggtcat gcagcattaa actttgacaa 1500

ggccactgac atcgtgacag gtctgaaaca gaaaacctcc tttgggagac 1550

caatgtggga caatgagttt tctacaatag ctacctccca ccccaagtct 1600

gtagtgggag ttttcttatg tggccctcgg actttggcaa agagcctgcg 1650

caaatgctgt caccgatatt ccagtctgga tcctagaaag gttcaattct 1700

acttcaacaa agaaaatttt tgagttatag gaataaggac ggtaatctgc 1750

attttgtctc tttgtatctt cagtaattta cttggtctcg tcaggtttga 1800

gcagtcactt taggaag 1817

<210> 54
<211> 1000
<212> DNA
<213> Homo sapiens

<400> 54
gggctccaaa ccacctcttg acaatgggaa actgggtggt taaccactgg 50

ttttcagttt tgtttctggt tgtttggtta gggctgaatg ttttcctgtt 100

tgtggatgcc ttcctgaaat atgagaaggc cgacaaatac tactacacaa 150

gaaaaatcct tgggtcttgg aagtggatcc ttgcaccggt cattctttat 200

atctgtgaaa ggatcctccg gttttaccgc tcccagcaga aggttgtgat 250

taccaaggtt gttatgcacc catccaaagt tttggaattg cagatgaaca 300

agcgtggctt cagcatggaa gtggggcagt atatctttgt taattgcccc 350

tcaatctctc tcctggaatg gcatcctttt actttgacct ctgctccaga 400

ggaagatttc ttctccattc atatccgagc agcaggggac tggacagaaa 450

atctcataag ggctttcgaa caacaatatt caccaattcc caggattgaa 500

gtggatggtc cctttggcac agccagtgag gatgttttcc agtatgaagt 550

ggctgtgctg gttggagcag gaattggggt cacccccttt gcttctatct 600

tgaaatccat ctggtacaaa ttccagtgtg cagaccacaa cctcaaaaca 650

aaaaaggttg gtcatgcagc attaaacttt gacaaggcca ctgacatcgt 700

gacaggtctg aaacagaaaa cctcctttgg gagaccaatg tgggacaatg 750

agttttctac aatagctacc tcccacccca agtctgtagt gggagtttc 800

ttatgtggcc ctcggacttt ggcaaagagc ctgcgcaaat gctgtcaccg 850

atattccagt ctggatccta gaaaggttca attctacttc aacaaagaaa 900

atttttgagt tataggaata aggacggtaa tctgcatttt gtctctttgt 950

atcttcagta atttacttgg tctcgtcagg tttgagcagt cactttagga 1000

<210> 55
<211> 1665
<212> DNA
<213> Homo sapiens

<400> 55
gggctccaaa ccacctcttg acaatgggaa actgggtggt taaccactgg 50

ttttcagttt tgtttctggt tgtttggtta gggctgaatg ttttcctgtt 100

tgtggatgcc ttcctgaaat atgagaaggc cgacaaatac tactacacaa 150

gaaaaatcct tgggtcaaca ttggcctgtg cccgagcgtc tgctctctgc 200

ttgaatttta acagcacgct gatcctgctt cctgtgtgtc gcaatctgct 250

gtccttcctg aggggcacct gctcattttg cagccgcaca ctgagaaagc 300

aattggatca caacctcacc ttccacaagc tggtggccta tatgatctgc 350

ctacatacag ctattcacat cattgcacac ctgtttaact ttgactgcta 400

tagcagaagc cgacaggcca cagatggctc ccttgcctcc attctctcca 450

gcctatctca tgatgagaaa aaggggggtt cttggctaaa tcccatccag 500

tcccgaaaca cgacagtgga gtatgtgaca ttcaccagca ttgctggtct 550

cactggagtg atcatgacaa tagccttgat tctcatggta acttcagcta 600

ctgagttcat ccggaggagt tattttgaag tcttctggta tactcaccac 650

cttttttatct tctatatcct tggcttaggg attcacggca ttggtggaat 700

tgtccggggt caaacagagg agagcatgaa tgagagtcat cctcgcaagt 750

gtgcagagtc ttttgagatg tgggatgatc gtgactccca ctgtaggcgc 800

cctaagtttg aagggcatcc ccctgagtct tggaagtgga tccttgcacc 850

ggtcattctt tatatctgtg aaaggatcct ccggttttac cgctcccagc 900

agaaggttgt gattaccaag gttgttatgc acccatccaa agttttggaa 950

ttgcagatga acaagcgtgg cttcagcatg gaagtggggc agtatatctt 1000

tgttaattgc ccctcaatct ctctcctgga atggcatcct tttactttga 1050

cctctgctcc agaggaagat ttcttctcca ttcatatccg agcagcaggg 1100

gactggacag aaaatctcat aagggctttc gaacaacaat attcaccaat 1150

tcccaggatt gaagtggatg gtccctttgg cacagccagt gaggatgttt 1200

tccagtatga agtggctgtg ctggttggag caggaattgg ggtcaccccc 1250

tttgcttcta tcttgaaatc catctggtac aaattccagt gtgcagacca 1300

caacctcaaa acaaaaaagg ttggtcatgc agcattaaac tttgacaagg 1350

ccactgacat cgtgacaggt ctgaaacaga aaacctcctt tgggagacca 1400

atgtgggaca atgagttttc tacaatagct acctcccacc ccaagtctgt 1450

agtgggagtt ttcttatgtg gccctcggac tttggcaaag agcctgcgca 1500

aatgctgtca ccgatattcc agtctggatc ctagaaaggt tcaattctac 1550

ttcaacaaag aaaatttttg agttatagga ataaggacgg taatctgcat 1600

tttgtctctt tgtatcttca gtaatttact tggtctcgtc aggtttgagc 1650

agtcacttta ggaag 1665

<210> 56
<211> 2400
<212> DNA
<213> Homo sapiens

<400> 56
ggactgcttt gtaactgcta agattgcaga cagaaatagc acacaaccac 50

tgtgagctgt atgcgattca gaaaccaaga ccaaattttg ctcactttca 100

ttaatcagtt gctcagatag aaggaaatga catctggttc tgtcttcttc 150

tacatcttaa tttttggaaa atatttttct catgggggtg gacaggatgt 200

caagtgctcc cttggctatt tcccctgtgg gaacatcaca aagtgcttgc 250

ctcagctcct gcactgtaac ggtgtggacg actgcgggaa tcaggccgat 300

gaggacaact gtggagacaa caatggatgg tccatgcaat ttgacaaata 350

ttttgccagt tactacaaaa tgacttccca atatcctttt gaggcagaaa 400

cacctgaatg tttggtcggt tctgtgccag tgcaatgtct ttgccaaggt 450

ctggagcttg actgtgatga aaccaattta cgagctgttc catcggtttc 500

ttcaaatgtg actgcaatgt cacttcagtg gaacttaata agaaagcttc 550

ctcctgattg cttcaagaat tatcatgatc ttcagaagct gtacctgcaa 600

aacaataaga ttacatccat ctccatctat gctttcagag gactgaatag 650

ccttactaaa ctgtatctca gtcataacag aataaccttc ctgaagccgg 700

gtgtttttga agatcttcac agactagaat ggctgataat tgaagataat 750

cacctcagtc gaatttcccc accaacattt tatggactaa attctcttat 800

tctcttagtc ctgatgaata acgtcctcac ccgtttacct gataaacctc 850

tctgtcaaca catgccaaga ctacattggc tggaccttga aggcaaccat 900

atccataatt taagaaattt gacttttatt tcctgcagta atttaactgt 950

tttagtgatg aggaaaaaca aaattaatca cttaaatgaa aatacttttg 1000

cacctctcca gaaactggat gaattggatt taggaagtaa taagattgaa 1050

aatcttccac cgcttatatt caaggacctg aaggagctgt cacaattgaa 1100

tctttcctat aatccaatcc agaaaattca agcaaaccaa tttgattatc 1150

ttgtcaaact caagtctctc agcctagaag ggattgaaat ttcaaatatc 1200

caacaaagga tgtttagacc tcttatgaat ctctctcaca tatattttaa 1250

gaaattccag tactgtgggt atgcaccaca tgttcgcagc tgtaaaccaa 1300

acactgatgg aatttcatct ctagagaatc tcttggcaag cattattcag 1350

agagtatttg tctgggttgt atctgcagtt acctgctttg gaaacatttt 1400

tgtcatttgc atgcgacctt atatcaggtc tgagaacaag ctgtatgcca 1450

tgtcaatcat ttctctctgc tgtgccgact gcttaatggg aatatattta 1500

ttcgtgatcg gaggctttga cctaaagttt cgtggagaat acaataagca 1550

tgcgcagctg tggatggaga gtactcattg tcagcttgta ggatctttgg 1600

ccattctgtc cacagaagta tcagttttac tgttaacatt tctgacattg 1650

gaaaaataca tctgcattgt ctatcctttt agatgtgtga gacctggaaa 1700

atgcagaaca attacagttc tgattctcat ttggattact ggttttatag 1750

tggctttcat tccattgagc aataaggaat ttttcaaaaa ctactatggc 1800

accaatggag tatgcttccc tcttcattca gaagatacag aaagtattgg 1850

agcccagatt tattcagtgg caatttttct tggtattaat ttggccgcat 1900

ttatcatcat agtttttcc tatggaagca tgttttatag tgttcatcaa 1950

agtgccataa cagcaactga aatacggaat caagttaaaa aagagatgat 2000

ccttgccaaa cgtttttct ttatagtatt tactgatgca ttatgctgga 2050

tacccatttt tgtagtgaaa tttctttcac tgcttcaggt agaaatacca 2100

ggtaccataa cctcttgggt agtgattttt attctgccca ttaacagtgc 2150

tttgaaccca attctctata ctctgaccac aagaccattt aaagaaatga 2200

ttcatcggtt ttggtataac tacagacaaa gaaaatctat ggacagcaaa 2250

ggtcagaaaa catatgctcc atcattcatc tgggtggaaa tgtggccact 2300

gcaggagatg ccacctgagt taatgaagcc ggacctttc acataccct 2350

gtgaaatgtc actgatttct caatcaacga gactcaattc ctattcatga 2400

<210> 57
<211> 2967
<212> DNA
<213> Homo sapiens

<400> 57
tagagatccc tcgacctcga cccacgcgtc cgaggaaaga aaaaaagagg 50

aatggaaaga gacagagaaa ggaaatggga gtggaaggag ggaggactgc 100

tttgtaactg ctaagattgc agacagaaat agcacacaac cactgtgagc 150

tgtatgcgat tcagaaacca agaccaaatt ttgctcactt tcattaatca 200

gttgctcaga tagaaggaaa tgacatctgg ttctgtcttc ttctacatct 250

taatttttgg aaaatatttt tctcatgggg gtggacagga tgtcaagtgc 300

tcccttggct atttcccctg tgggaacatc acaaagtgct tgcctcagct 350

cctgcactgt aacggtgtgg acgactgcgg gaatcaggcc gatgaggaca 400

actgtggaga caacaatgga tggtctctgc aatttgacaa atattttgcc 450

agttactaca aaatgacttc ccaatatcct tttgaggcag aaacacctga 500

atgtttggtc ggttctgtgc cagtgcaatg tctttgccaa ggtctggagc 550

ttgactgtga tgaaaccaat ttacgagctg ttccatcggt ttcttcaaat 600

gtgactgcaa tgtcacttca gtggaactta ataagaaagc ttcctcctga 650

ttgcttcaag aattatcatg atcttcagaa gctgtacctg caaaacaata 700

agattacatc catctccatc tatgctttca gaggactgaa tagccttact 750

aaactgtatc tcagtcataa cagaataacc ttcctgaagc cgggtgtttt 800

tgaagatctt cacagactag aatggctgat aattgaagat aatcacctca 850

gtcgaatttc cccaccaaca ttttatggac taaattctct tattctctta 900

gtcctgatga ataacgtcct cacccgttta cctgataaac ctctctgtca 950

acacatgcca agactacatt ggctggacct tgaaggcaac catatccata 1000

atttaagaaa tttgactttt atttcctgca gtaatttaac tgttttagtg 1050

atgaggaaaa acaaaattaa tcacttaaat gaaaatactt ttgcacctct 1100

ccagaaactg gatgaattgg atttaggaag taataagatt gaaaatcttc 1150

caccgcttat attcaaggac ctgaaggagc tgtcacaatt gaatctttcc 1200

tataatccaa tccagaaaat tcaagcaaac caatttgatt atcttgtcaa 1250

actcaagtct ctcagcctag aagggattga aatttcaaat atccaacaaa 1300

ggatgtttag acctcttatg aatctctctc acatatattt taagaaattc 1350

cagtactgtg ggtatgcacc acatgttcgc agctgtaaac caaacactga 1400

tggaatttca tctctagaga atctcttggc aagcattatt cagagagtat 1450

ttgtctgggt tgtatctgca gttacctgct ttggaaacat ttttgtcatt 1500

tgcatgcgac cttatatcag gtctgagaac aagctgtatg ccatgtcaat 1550

catttctctc tgctgtgccg actgcttaat gggaatatat ttattcgtga 1600

tcggaggctt tgacctaaag tttcgtggag aatacaataa gcatgcgcag 1650

ctgtggatgg agagtactca ttgtcagctt gtaggatctt tggccattct 1700

gtccacagaa gtatcagttt tactgttaac atttctgaca ttggaaaaat 1750

acatctgcat tgtctatcct tttagatgtg tgagacctgg aaaatgcaga 1800

acaattacag ttctgattct catttggatt actggtttta tagtggcttt 1850

cattccattg agcaataagg aatttttcaa aaactactat ggcaccaatg 1900

gagtatgctt ccctcttcat tcagaagata cagaaagtat tggagcccag 1950

atttattcag tggcaatttt tcttggtatt aatttggccg catttatcat 2000

catagttttt tcctatggaa gcatgtttta tagtgttcat caaagtgcca 2050

taacagcaac tgaaatacgg aatcaagtta aaaaagagat gatccttgcc 2100

aaacgttttt tctttatagt atttactgat gcattatgct ggatacccat 2150

ttttgtagtg aaatttcttt cactgcttca ggtagaaata ccaggtacca 2200

taacctcttg ggtagtgatt tttattctgc ccattaacag tgctttgaac 2250

ccaattctct atactctgac cacaagacca tttaaagaaa tgattcatcg 2300

gttttggtat aactacagac aaagaaaatc tatggacagc aaaggtcaga 2350

aaacatatgc tccatcattc atctgggtgg aaatgtggcc actgcaggag 2400

atgccacctg agttaatgaa gccggacctt ttcacatacc cctgtgaaat 2450

gtcactgatt tctcaatcaa cgagactcaa ttcctattca tgactgactc 2500

tgaaattcat ttcttcgcag agaatactgt gggggtgctt catgagggat 2550

ttactggtat gaaatgaata ccacaaaatt aatttataat aatagctaag 2600

ataaatattt tacaaggaca tgaggaaaaa taaaaatgac taatgctctt 2650

acaaagggaa gtaattatat caataatgta tatatattag tagacatttt 2700

gcataagaaa ttaagagaaa tctacttcag taacattcat tcatttttct 2750

aacatgcatt tattgagtac ccactactat gtgcatagca ttgcaatata 2800

gtcctggaag tagacagtgc agaacctttc aatctgtaga tggtgtttaa 2850

tgacaaaaga ctatacaaag tccatctgca gttcctagtt taaagtagag 2900

ctttacctgt catgtgcatc agcaagaatc ataggcactt ttaaataaag 2950

gtttaaagtt ttggaat 2967

<210> 58
<211> 3032
<212> DNA
<213> Homo sapiens

<400> 58
gatccattgt gctggcgcgg attctttatc actggaagga ggaggagggg 50

gagagtcgct cccgccgggc gagcatgggg cgcctggcct cgaggccgct 100

gctgctggcg ctcctgtcgt tggctctttg ccgagggcgt gtggtgagag 150

tccccacagc gaccctggtt cgagtggtgg gcactgagct ggtcatcccc 200

tgcaacgtca gtgactatga tggccccagc gagcaaaact ttgactggag 250

cttctcatct ttggggagca gctttgtgga gcttgcaagc acctgggagg 300

tggggttccc agcccagctg taccaggagc ggctgcagag gggcgagatc 350

ctgttaaggc ggactgccaa cgacgccgtg gagctccaca taaagaacgt 400

ccagccttca gaccaaggcc actacaaatg ttcaaccccc agcacagatg 450

ccactgtcca gggaaactat gaggacacag tgcaggttaa agtgctggcc 500

gactccctgc acgtgggccc cagcgcgcgg cccccgccga gcctgagcct 550

gcgggagggg gagcccttcg agctgcgctg caccgccgcc tccgcctcgc 600

cgctgcacac gcacctggcg ctgctgtggg aggtgcaccg cggcccggcc 650

aggcggagcg tcctcgccct gacccacgag ggcaggttcc acccgggcct 700

ggggtacgag cagcgctacc acagtgggga cgtgcgcctc gacaccgtgg 750

gcagcgacgc ctaccgcctc tcagtgtccc gggctctgtc tgccgaccag 800

ggctcctaca ggtgtatcgt cagcgagtgg atcgccgagc agggcaactg 850

gcaggaaatc caagaaaagg ccgtggaagt tgccaccgtg gtgatccagc 900

cgacagttct gcgagcagct gtgcccaaga atgtgtctgt ggctgaagga 950

aaggaactgg acctgacctg taacatcaca acagaccgag ccgatgacgt 1000

ccggcccgag gtgacgtggt ccttcagcag gatgcctgac agcaccctac 1050

ctggctcccg cgtgttggcg cggcttgacc gtgattccct ggtgcacagc 1100

tcgcctcatg ttgctttgag tcatgtggat gcacgctcct accatttact 1150

ggttcgggat gttagcaaag aaaactctgg ctactattac tgccacgtgt 1200

ccctgtgggc acccggacac aacaggagct ggcacaaagt ggcagaggcc 1250

gtgtcttccc cagctggtgt gggtgtgacc tggctagaac cagactacca 1300

ggtgtacctg aatgcttcca aggtccccgg gtttgcggat gaccccacag 1350

agctggcatg ccgggtggtg gacacgaaga gtggggaggc gaatgtccga 1400

ttcacggttt cgtggtacta caggatgaac cggcgcagcg acaatgtggt 1450

gaccagcgag ctgcttgcag tcatggacgg ggactggacg ctaaaatatg 1500

gagagaggag caagcagcgg gcccaggatg gagactttat tttttctaag 1550

gaacatacag acacgttcaa tttccggatc caaaggacta cagaggaaga 1600

cagaggcaat tattactgtg ttgtgtctgc ctggaccaaa cagcggaaca 1650

acagctgggt gaaaagcaag gatgtcttct ccaagcctgt taacatattt 1700

tgggcattag aagattccgt gcttgtggtg aaggcgaggc agccaaagcc 1750

tttctttgct gccggaaata catttgagat gacttgcaaa gtatcttcca 1800

agaatattaa gtcgccacgc tactctgttc tcatcatggc tgagaagcct 1850

gtcggcgacc tctccagtcc caatgaaacg aagtacatca tctctctgga 1900

ccaggattct gtggtgaagc tggagaattg gacagatgca tcacgggtgg 1950

atggcgttgt tttagaaaaa gtgcaggagg atgagttccg ctatcgaatg 2000

taccagactc aggtctcaga cgcagggctg taccgctgca tggtgacagc 2050

ctggtctcct gtcaggggca gcctttggcg agaagcagca accagtctct 2100

ccaatcctat tgagatagac ttccaaacct caggtcctat atttaatgct 2150

tctgtgcatt cagacacacc atcagtaatt cggggagatc tgatcaaatt 2200

gttctgtatc atcactgtcg agggagcagc actggatcca gatgacatgg 2250

cctttgatgt gtcctggttt gcggtgcact cttttggcct ggacaaggct 2300

cctgtgctcc tgtcttccct ggatcggaag ggcatcgtga ccacctcccg 2350

gagggactgg aagagcgacc tcagcctgga gcgcgtgagt gtgctggaat 2400

tcttgctgca agtgcatggc tccgaggacc aggactttgg caactactac 2450

tgttccgtga ctccatgggt gaagtcacca acaggttcct ggcagaagga 2500

ggcagagatc cactccaagc ccgtttttat aactgtgaag atggatgtgc 2550

tgaacgcctt caagtatccc ttgctgatcg gcatcggtct gtccacggtc 2600

atcgggctcc tgtcctgtct catcgggtac tgcagctccc actggtgttg 2650

taagaaggag gttcaggaga cacggcgcga gcgccgcagg ctcatgtcga 2700

tggagatgga ctaggctggc ccgggagggg agtgacagag ggacgttcta 2750

ggagcaattg gggcaagaag aggacagtga tattttaaaa caaagtgtgt 2800

tacactaaaa accagtcctc tctaatctca ggtgggactt ggcgctctct 2850

cttttctgca tgtcaagttc tgagcgcgga catgtttacc agcacacggc 2900

tcttcttccc acggcacttt ctgatgtaac aatcgagtgt gtgttttccc 2950

aactgccagt gataaagaat ccgcgccagc acaatggatc tcgagggatc 3000

ttccatacct accagttctg cgcctgcagg tc 3032

<210> 59
<211> 2771
<212> DNA
<213> Homo sapiens

<400> 59
cccacgcgtc cgcggtaaat cacttcttgg aggtgccctg cacgccggtc 50

ctgggagcag gcggcctccc gggggtgcgg gagccccact cctccgtggt 100

gtgttccatt tgcttcccac atctggagga gctgacgtgc cagcctcccc 150

cagcaccacc cagggacggg aggcatgagc cggtcaaggc acctgggcaa 200

aatccggaag cgtctggaag atgtcaagag ccagtgggtc cggccagcca 250

gggctgactt tagtgacaac gagagtgccc ggctggccac ggacgccctc 300

ttggatgggg gttctgaagc ctactggcgg gtgctcagcc aggaaggcga 350

ggtggacttc ttgtcctcgg tggaggccca gtacatccag gcccaggcca 400

gggagccccc gtgtccccca gacaccctgg gaggggcgga agcaggccct 450

aagggactgg actccagctc cctacagtcc ggcacctact tccctgtggc 500

ctcagagggc agcgagccgg ccctactgca cagctgggcc tcagctgaga 550

agccctacct gaaggaaaaa tccagcgcca ctgtgtactt ccagaccgtc 600

aagcacaaca acatcagaga cctcgtccgc cgctgcatca cccggactag 650

ccaggtcctg gtcatcctga tggatgtgtt cacggatgtg gagatcttct 700

gtgacattct agaggcagcc aacaagcgtg gggtgttcgt ttgtgtgctc 750

ctggaccagg gaggtgtgaa gctcttccag gagatgtgtg acaaagtcca 800

gatctctgac agtcacctca agaacatttc catccggagt gtggaaggag 850

agatatactg tgccaagtca ggcaggaaat tcgctggcca aatccgggag 900

aagttcatca tctcggactg gagatttgtc ctgtctggat cttacagctt 950

cacctggctc tgcggacacg tgcaccggaa catcctctcc aagttcacag 1000

gccaggcggt ggagctgttt gacgaggagt tccgccacct ctacgcctcc 1050

tccaagcctg tgatgggcct gaagtccccg cggctggtcg cccccgtccc 1100

gcccggagca gccccggcca atggccgcct tagcagcagc agtggctccg 1150

ccagtgaccg cacgtcctcc aacccccttca gcggccgctc ggcaggcagc 1200

caccccggta cccgaagtgt gtccgcgtct tcagggccct gtagccccgc 1250

ggccccacac ccgcctccac cgccccggtt ccagccccac caaggccctt 1300

ggggagcccc gagtccccag gcccacctct ccccgcggcc ccacgacggc 1350

ccgcccgccg ctgtctacag caacctgggg gcctacaggc ccacgcggct 1400

gcagctggag cagctgggcc tggtgccgag gctgactcca acctggaggc 1450

ccttcctgca ggcctcccct cacttctgaa ggtcccatcc cctgctgccc 1500

tccgcaggcc cagggctggg cactccctga gacccaaaga cccacctcaa 1550

cgacgagtgg cgttgagcca cttccctttg aaaagacact caaaatcact 1600

gccatggttc aatgttccca ggccccaggc catccacttg ccggccccca 1650

ccagttcttg ggttccccgc tctagtttga cctgtgcagc acattccaga 1700

aggttccagg ggaggttgtg gggcagctag aggacaaaat catgaaaaca 1750

gagtccctgt cttccagaga tcatccgggg ctttaatatt aatggccccc 1800

aaaactccgt aagaagcagg aaatgcagcc caagttttac aaatgggtaa 1850

acagaggcac tgagagatag atggtagttt ggtacttctg gttcccagtg 1900

cccaggaatg gtccactccc aagaaattca ggaaagaaag actgaggaga 1950

aggtgtggga acattctgga tgtttcggga gagttgggga aactcctcct 2000

cttaggaaag gctaatacta gggtatcctt gggcccaatg aattaggggt 2050

gaggccccag aacccgttat ctatgagttg tatgggggag ccatctgaag 2100

ctgtagccac cagggatgca gctagctgag gagtttgggg tgttgggttg 2150

gacaaggcag gttagtagac tcagattctt gcttcaaaga gccttgggct 2200

ggcctggagg tccctggagt ctagactgga cctaggagct tgagttgtca 2250

ggggccagga ctggccccac tgcagtgccc aggccagtct tgagcagcag 2300

ggagggctca gctgtcccca gatccaggtg cctctgacca gcctggtcac 2350

ctcctgagga ataaatgctg aacctcacaa gccccatcat tcatttcttc 2400

tcaattcaca gtgcccctct ttgtttctgg ggtggaacta ggtcctgagg 2450

gcacagccta gctgagtgca aagaaatata ggatgcttag aaagcataca 2500

ggaggggcca ggcgtggtgg ctcatgcctg taatcccaga actttgggat 2550

gccaaggtgg ttggattacc tgagatcagg tggattacct ggtctcgaga 2600

ccagcctgac caatatggtg aaaccccgtc tctactaaaa atacaaaaat 2650

taggctgaga caggagaatt gcttgaaccc aggaagcaga ggttgcaatg 2700

agctgagatt gcatcactgc actccagcat gggcaacaaa gcaagactcc 2750

gtcacagaaa aaaaaaaaaa a 2771

<210> 60
<211> 4060
<212> DNA
<213> Homo sapiens

<400> 60
atggctctgc cgcagccgcc gccgcccgcc ggggcgcgtg acccggtgac 50

cccggagcac tgggcatccg gacccccctc gcaggcgcag ccgctattaa 100

ggcaggaggc taaagaggag gaggaagggg aggagacagg cgtccagggc 150

gcctggggaa ccggcacggc ggagcagcgg cggcggggct ggggggaggc 200

cgccgagtcg gccgcggccg aggaggggca ggcggaggtc gggggcgccg 250

cggccgcggg gtcaggttcc ccggcgggag gcgcaggtgg cggcctgggg 300

agctggcggc cgctcctcgc ctggctgcag cggcggcagc cccagtgctg 350

cccctgtgcg gcgccccttt cccgctccgc cgcgcactgt tgtcatggag 400

gaaccaagat ggcggctctg gcctacaacc tgggcaagcg ggagatcaac 450

cactacttca gcgtgaggag cgccaaggtg ctggcgctgg tggccgtgct 500

gctgctcgca gcgtgccacc tcgcctcccg ccgctaccga ggcaatgatt 550

cgtgtgaata ccttctctca agtggcagat ttcttggaga gaaagtttgg 600

caacctcaca gttgtatgat gcataaatac aaaatcagtg aagcaaagaa 650

ctgccttgta gataaacata ttgcatttat tggagattcc agaattcgtc 700

aattgtttta ttcttttgta aaaataatta atccccaatt caaagaagaa 750

ggaaataagc atgaaaacat tccttttgaa gacaagactg catcagttaa 800

agtggatttt ctgtggcatc ctgaagttaa tggttctatg aaacagtgta 850

tcaaagtgtg gactgaggat tccattgcaa agccacatgt gattgtagca 900

ggagctgcca catggtccat caagattcac aatggtagca gtgaagcgct 950

ttctcaatat aaaatgaaca tcacctccat agcaccactt ttagaaaaat 1000

tggcaaagac tagtgatgtt tattgggtct tacaagatcc tgtttatgaa 1050

gatctattaa gtgaaaatag gaagatgatc actaatgaga agatagatgc 1100

ttacaatgaa gctgcagtca gtattttgaa tagtagcacc agaaattcta 1150

aatcaaatgt taagatgttc agtgtttcca aattaattgc tcaagaaacc 1200

atcatggaat ctttggatgg cttacatctt cctgaatcga gcagagaaac 1250

tactgcaatg attcttatga atgtgtattg caataagatt ttgaagcctg 1300

tagatgggtc ctgttgtcaa cctcggcctc ctgttactct catacagaag 1350

ctagctgctt gttttttcac tttatctatt atcggatatt taatttttta 1400

cataattcat cgtaatgctc atcggaagaa taagccgtgt actgatttgg 1450

aaagtggaga ggaaaagaaa aatattatca ataccccctgt gtcttcatta 1500

gaaatacttt tacaatcttt ctgcaaactt ggcctgatta tggcatattt 1550

ctatatgtgt gaccgtgcaa atctgttcat gaaggaaaac aaattttata 1600

cacattcatc tttctttatt ccaattatct acattttggt tttgggagta 1650

ttttataatg aaaatactaa agagactaaa gtattaaata gagaacaaac 1700

agacgaatgg aaaggctgga tgcaacttgt gattttgatt tatcacattt 1750

ctggagcaag tacatttttg cctgtataca tgcacattcg agttctggtt 1800

gctgcatatt tatttcagac agggtatggg catttctcat acttttggat 1850

aaaaggagat tttggaatct atagagtatg tcaggtttta tttcgtctca 1900

atttcctggt agtggtgtta tgtatagtaa tggatcgacc ttatcaattc 1950

tattactttg tccccttggt cactgtatgg ttcatggtca tatatgttac 2000

tttagcacta tggccacaaa taatccaaaa aaaagcaaac ggaaattgtt 2050

tctggcattt tggcttactg ttgaaactag gctttttgct gttattcata 2100

tgtttttgg catattctca gggtgcattt gagaagatct tttctctttg 2150

gccattgtcc aagtgttttg aactgaaagg gaatgtatat gaatggtggt 2200

tcagatggag gttagaccgt tatgtagttt tccacggaat gctgtttgct 2250

tttatttatc tggctttgca gaagcgtcaa atactttctg aaggaaaggg 2300

tgaacctctt ttttcaaaca aaatttcaaa ttttctgttg tttatttcag 2350

tagtttcttt cttgacctat tccatctggg ctagcagttg taaaaacaaa 2400

gcagagtgca atgaactcca tccgtctgtt tctgtggtac agattttagc 2450

cttcatccta ataagaaaca tccctggata tgcccgttca gtttacagtt 2500

catttttgc ttggtttgga aaaatttcat tagagctatt tatttgccag 2550

tatcacatat ggctggcagc ggacacaagg ggtatcttgg tactgatacc 2600

tggaaaccct atgctcaaca tcattgtcag cactttcata tttgtttgtg 2650

tggcacatga aatttctcag atcactaatg atcttgcaca gattattatt 2700

cctaaagata actcatctct cttgaaaagg ttggcatgta tagctgcatt 2750

tttttgtgga ctcctcatct tatcatccat tcaagataaa tcaaaacatt 2800

aggttccaaa aattctaaaa aacctaaact cttcaggcta cctttgtgtg 2850

tctctagaag agaaaagcat ctatctggag atataaatgt gtwtgtaaat 2900

ataaacgttt gtggcaagag gacagttctg tgacatctgt tgaacatatg 2950

tggttgtata tattggaaat gtacatatcc aatatgaaat actaaaacaa 3000

acaaacaaac aaaaaaccag aatgcattgt ataggattgc atgtgaagtc 3050

ttttctactg aatctatatt tccatttgta agtgatttta agttaacata 3100

tgaaggcagg gaaatgatta cctttccagt aaaaagtata gataatttaa 3150

ttaacttagt gacaccacca agtgttttga tataactaaa tttgtggtaa 3200

taagactgtc tgcacctgta ttcattgtgg aacttcctct ttcattggaa 3250

actttcttac tcaagaatga cggcagtatt gttttcttat atgtgcaatg 3300

aagtggaatg ataaacagta tgcctttaat ttatatgtgt tcttgttctg 3350

atgttgtttc ctgaaatgat ttttcttcct aactgtggtt ttcgggtatg 3400

caagcctaaa tctttgtaca ctttgtctca cagaatagtt ctgaggctcc 3450

atgacagggt tttgtcattg ttgatgttat tgttgcttcg ttttataaaa 3500

aagccaaaat tttttttcca atccaaacgt tcacctgttt cctttcctca 3550

agctatacca gtgtaatacc agttaccctg tggatccatt taatatgtta 3600

tccccactaa ttaattttcg tatattattt ccaatatttg gaaagctctt 3650

tatagccatt tggtatttcc tattacccac ctcctatttt aaatatttat 3700

cagtctaaac ttgtgcagtg tagtaaacat gcaagttgtt acgattgagc 3750

tgtattacca taagtagaat tttaagtaaa ctggtgaatt tgggcaataa 3800

atgtttttgc tttttgtttg attttttttt acaagctaac tgttagaggt 3850

atacatttat ttatctgttg tacagatttg attatgattt taatgtttga 3900

aagattgcac ttgtttgctt ttactatatg tggggtaaaa tatattttct 3950

gttcacagta tatgaaaata tggagtaatt taaacagtaa ataaacattc 4000

tgtggatgct tatttttgta ttggcaaagt atcaattaaa ctatatgtgt 4050

tctttttcaa 4060

<210> 61
<211> 5472
<212> DNA
<213> Homo sapiens

<400> 61
gaaagacata cacacttcat gtaatgctac ctgcaagtct ccctagaaaa 50

gcagtttttg taggtgaaaa caatgaagcc aggtaatatt gcaaggaggc 100

tgtaatttta gcagacctac caacaacact gatgtaggaa gctcattatt 150

ttaatttctg gagcctttta atttttttctt tagaaagtgt ataaataatt 200

gcagtgctgc tttgcttcca aaactgggca gtgagttcaa caacaacgac 250

aacaacagcc gcagctcatc ctggccgtca tggagtttct tgaaagaacg 300

tatcttgtga atgataaagc tgccaagatg tatgctttca cactagaaag 350

aaggagctgc aaatgaacac ttcatagcaa tgtggaactc caacagaaac 400

cggtgaataa agatcagtgt cccagagaga gaccagagga gctggagtca 450

ggaggcatgt accactgcca cagtggctcc aagcccacag aaaagggggc 500

gaatgagtac gcctatgcca agtggaaact ctgttctgct tcagcaatat 550

gcttcatttt catgattgca gaggtcgtgg gtgggcacat tgctgggagt 600

cttgctgttg tcacagatgc tgcccacctc ttaattgacc tgaccagttt 650

cctgctcagt ctcttctccc tgtggttgtc atcgaagcct ccctctaagc 700

ggctgacatt tggatggcac cgagcagaga tccttggtgc cctgctctcc 750

atcctgtgca tctgggtggt gactggcgtg ctagtgtacc tggcatgtga 800

gcgcctgctg tatcctgatt accagatcca ggcgactgtg atgatcatcg 850

tttccagctg cgcagtggcg gccaacattg tactaactgt ggttttgcac 900

cagagatgcc ttggccacaa tcacaaggaa gtacaagcca atgccagcgt 950

cagagctgct tttgtgcatg cccttggaga tctatttcag agtatcagtg 1000

tgctaattag tgcacttatt atctacttta agccagagta taaaatagcc 1050

gacccaatct gcacattcat cttttccatc ctggtcttgg ccagcaccat 1100

cactatctta aaggacttct ccatcttact catggaaggt gtgccaaaga 1150

gcctgaatta cagtggtgtg aaagagctta ttttagcagt cgacggggtg 1200

ctgtctgtgc acagcctgca catctggtct ctaacaatga atcaagtaat 1250

tctctcagct catgttgcta cagcagccag ccgggacagc caagtggttc 1300

ggagagaaat tgctaaagcc cttagcaaaa gctttacgat gcactcactc 1350

accattcaga tggaatctcc agttgaccag gaccccgact gccttttctg 1400

tgaagacccc tgtgactagc tcagtcacac cgtcagtttc ccaaatttga 1450

caggccacct tcaaacatgc tgctatgcag tttctgcatc atagaaaata 1500

aggaaccaaa ggaagaaatt catgtcatgg tgcaatgcac attttatcta 1550

tttatttagt tccattcacc atgaaggaag aggcactgag atccatcaat 1600

caattggatt atatactgat cagtagctgt gttcaattgc aggaatgtgt 1650

atatagatta ttcctgagtg gagccgaagt aacagctgtt tgtaactatc 1700

ggcaatacca aattcatctc ccttccaata atgcatcttg agaacacata 1750

ggtaaatttg aactcaggaa agtcttacta gaaatcagtg gaagggacaa 1800

atagtcacaa aattttacca aaacattaga aacaaaaaat aaggagagcc 1850

aagtcaggaa taaaagtgac tctgtatgct aacgccacat tagaacttgg 1900

ttctctcacc aagctgtaat gtgatttttt tttctactct gaattggaaa 1950

tatgtatgaa tatacagaga agtgcttaca actaattttt atttacttgt 2000

cacattttgg caataaatcc ctcttatttc taaattctaa cttgtttatt 2050

tcaaaacttt atataatcac tgttcaaaag gaaatatttt cacctaccag 2100

agtgcttaaa cactggcacc agccaaagaa tgtggttgta gagacccaga 2150

agtcttcaag aacagccgac aaaaacattc gagttgaccc caccaagttg 2200

ttgccacaga taatttagat atttacctgc aagaaggaat aaagcagatg 2250

caaccaattc attcagtcca cgagcatgat gtgagcactg ctttgtgcta 2300

gacattgggc ttagcattga aactataaag aggaatcaga cgcagcaagt 2350

gcttctgtgt tctggtagca actcaacact atctgtggag agtaaactga 2400

agatgtgcag gccaacattc tggaaatcct atgtcactgg gtttggtttg 2450

gaaacctgga cttctgcatt tttaaaagtt acccagagat gcttctaaag 2500

atgagccata gtctagaaga ttgtcaacca caggagttca ttgagtggga 2550

cagctagaya catacattgg cagctacaat agtatcatga attgcaatga 2600

tgtagtgggg tataaaagga aagcgatgga tattgccgga tgggcatggc 2650

cagtgatgtt tcacgtcatt gaggtgacag ctctgctgga ctttgaatta 2700

catatggagg ctctccagga agacgaagaa gagaaggaca ttctaggcaa 2750

aaagaagact aggcacaagg cacacttatg tttgtctgtt agcttttagt 2800

tgaaaaagca aaatacatga tgcaaagaaa cctctccacg ctgtgatttt 2850

taaaactaca tacttttgc aactttatgg ttatgagtat tgtagagaac 2900

aggagatagg tcttagatga tttttatgtt gttgtcagac tctagcaagg 2950

tactagaaac ctagcaggca ttaataattg ttgaggcaat gactctgagg 3000

ctatatctgg gccttgtcat tatttatcat ttatatttgt atttttttct 3050

gaaatttgag ggccaagaaa acattgactt tgactgagga ggtcacatct 3100

gtgccatctc tgcaaatcaa tcagcaccac tgaaataact acttagcatt 3150

ctgctgagct ttccctgctc agtagagaca aatatactca tcccccacct 3200

cagtgagctt gtttaggcaa ccaggattag agctgctcag gttcccaacg 3250

tctcctgcca catcgggttc tcaaaatgga aagaatggtt tatgccaaat 3300

cacttttcct gtctgaagga ccactgaatg gttttgtttt tccatatttt 3350

gcataggacg ccctaaagac taggtgactt ggcaaacaca caagtgttag 3400

tataattctt tgcttctgct tctttttgaa aatcatgttt agatttgatt 3450

ttaagtcaga aattcactga atgtcaggta atcattatgg agggagattt 3500

gtgtgtcaac caaagtaatt gtcccatggc cccagggtat ttctgttgtt 3550

tccctgaaat tctgcttttt tagtcagcta gattgaaaac tctgaacagt 3600

agatgtttat atggcaaaat gcaagacaat ctataaggga gattttaagg 3650

attttgagat gaaaaaacag atgctactca ggggctttat ggaccatcca 3700

tcaattctga agttctgact ctcccattac cctttccctg gtgtggtcag 3750

aactccaggt cactggaagt tagtggaatc atgtagttga attctttact 3800

tcaagacatt gtattctctc cagctatcaa aacattaatg atcttttatg 3850

tctttttttt gttattgtta tactttaagt tctggggtac atgtgcggaa 3900

catgtaggtt tgttacatag gtatacatgt gccatggtgg tttgctgcac 3950

tcatcaacct gtcatctaca ttcttttatg tctgtctttc aaagcaacac 4000

tctgttcttc tgagtagtga aatcaggtca actttaccac cagcctccat 4050

ttttaatatg cttcaccatc atccagcacc tacttaagat ttatctaggg 4100

ctctgtggtg atgttaggac ccataaaaga aatttatgcc ttccatatgt 4150

ttggttacag atgggaaatg ggaatgttga aggacatgaa agaaaggatg 4200

tttacacatt aagcatcagt tctgaagcta gattgtctga gtttgaatct 4250

tagctcttcc ctttattagc tctgtgacct cgagctagtt acttaaatgc 4300

tctgatcctc tatttcctga tcagtgaaac ctccctattc aaatgtgtga 4350

gagtttaata aattaggaca cttaaaaatg ttggagcagt gcatagcatg 4400

tagtgttcag tacatgttaa atgttgtttt ttattatgta caaacatgtg 4450

tgggcacaga attttaaatc atctcaactt ttgagaaatt ttgagttatc 4500

aacaccgttc ccacaagaca gtggcaaaat tattggtgag aattaaacag 4550

ctgtttctca gaggaagcaa tggaggcttg ctgggataaa ggcatttact 4600

gagaggctgt tacctagtga gagtgatgaa ttaattaaaa tagtcgaatc 4650

cctttctgac tgtctctgaa agcttccgct tttatctttg aagagcagaa 4700

ttgtcacccc aaggacattt attaataaaa agaacaactg tccagtgcaa 4750

tgaaggcaaa gtcataggtc tcccaagtct taccccattc ctgtgaaata 4800

tcaagttctt ggcttttctc tgtcatgtag cctcaacttt ctccgaccgg 4850

gtgcatttct ttctctggtt tctaaattgc cagtggcaaa tttggatcac 4900

ttacttaata tctgttaaat tttgtgaccc aacaaagtct tttagcactg 4950

tggtgtcaaa aagaaaaaca cctcccaggc atatacattt tatagattcc 5000

tggagaatgt tgctctccag ctccatcccc acccaatgaa atatgatcca 5050

gagagtcttg caaagagaca agcctcattt tccacaatta gctctaaagt 5100

gcctccagga aatgattttc tcagctcatc tctctgtatt ccctgttttg 5150

gatcacaggg caatctgttt aaatgactaa ttacagaaat cattaaaggc 5200

accaagcaaa tgtcatctct gaatacacac atcccaagct ttacaaatcc 5250

tgcctggctt gacagtgatg aggccactta acagtccagc gcaggcggat 5300

gttaaaaaaa ataaaaaggt gaccatctgc ggtttagttt tttaactttc 5350

tgatttcaca cttaacgtct gtcattctgt tactgggcac ctgtttaaat 5400

tctattttaa aatgttaatg agtgttgttt aaaataaaat caggaaagag 5450

agaaaaaaaa aaaaaaaaaa aa 5472

<210> 62
<211> 5170
<212> DNA

<213> Homo sapiens

<400> 62
gcggcgggag agctgctggc tcgcccggat cccgggagct gcctggaggc 50

gggcccggcc cggggaaggt gagcggctgc gggacccagc ccctcgccgg 100

gagcgggcac catggtgctg tcggtgcctg tgatcgcgct gggcgccacg 150

ctgggcacag ccaccagcat cctcgcgttg tgcggggtca cctgcctgtg 200

tcggcacatg caccccaaga aggggctgct gccgcgggac caggaccccg 250

acctggagaa ggcgaagccc agcttgctcg ggtctgcaca acagttcaat 300

gttaaaaagt ccacggaacc tgttcagccc cgtgccctcc tcaagttccc 350

agacatctat ggacccaggc cagctgtgac ggctccagag gtcatcaact 400

atgcagacta ttcactgagg tctacggagg agcccactgc acctgccagc 450

ccccaacccc cgaatgacag tcgcctcaag aggcaggtca cagaggagct 500

gttcatcctc cctcagaatg gtgtggtgga ggatgtctgt gtcatggaga 550

cctggaaccc agagaaggct gccagttgga accaggcccc caaactccac 600

tactgcctgg actatgactg tcagaaggca gaattgtttg tgactcgcct 650

ggaagctgtg accagcaacc acgacggagg ctgtgactgc tacgtccaag 700

ggagtgtggc caataggacc ggctctgtgg aggctcagac agccctaaag 750

aagcggcagc tgcacaccac ctgggaggag ggcctggtgc tccccctggc 800

ggaggaggag ctccccacag ccaccctgac gctgaccttg aggacctgcg 850

accgcttctc ccgtcacagc gtggccgggg agctccgcct gggcctggac 900

gggacatctg tgcctctagg ggctgcccag tggggcgagc tgaagacttc 950

agcgaaggag ccatctgcag gagctggaga ggtcctacta tccatcagct 1000

acctcccggc tgccaaccgc ctcctggtgg tgctgattaa agccaagaac 1050

ctccactcta accagtccaa ggagctcctg gggaaggatg tctctgtcaa 1100

ggtgaccttg aagcaccagg ctcggaagct gaagaagaag cagactaaac 1150

gagctaagca caagatcaac cccgtgtgga acgagatgat catgtttgag 1200

ctgcctgacg acctgctgca ggcctccagt gtggagctgg aagtgctggg 1250

ccaggacgat tcagggcaga gctgtgcgct tggccactgc agcctgggcc 1300

tgcacacctc gggctctgag cgcagccact gggaggagat gctcaaaaac 1350

cctcgccggc agattgccat gtggcaccag ctgcacctgt aaccagctgc 1400

ccagctgcct cccttcttgg acagccctga cccgtcctct gcaacctcct 1450

ttctgtgccc ctttcctcat tctgacaccc agaagacagt gacagatgtg 1500

tttgcaaggc tgggatggct ctctcatcat actcttgttt cttagaaata 1550

agcaagacag agcaggaaat ggaatatgcg ggtcacactg aggaatgcat 1600

tttgctcatc tgtgttattg aaggaggtgc ttattaaata cagttcctat 1650

gcctgtttta taggtggggt taggccagat gcagagaaag ctaaatgtgg 1700

gaatcatgga tgcaaagaag aatttggctt tttgaaaaac aagcatttca 1750

aaaatgatga aggaagtgaa agtatcctgg atcaactcct agagttagag 1800

attgcccagg tggaaagaaa ccttagccag cgttcaatca agctcaccat 1850

gcagggcagt cacccggcag ttctcaaact ttagcatgtg aagagtcacc 1900

agcagattcc tgggctcgcc tggagacatt cctagtcggt attcctggtc 1950

gaagcccagg agccttcctt tttaacaagc tgatgtagag ggtggagcac 2000

tgtatgtgga gaaattcctt ctacaatatt ccacacaggt ttttggccac 2050

agtccttgat ggagtcccaa aaccatggtg cagccagttc caatgctgga 2100

cacctcaacc atcagggtga aatctggggc ctcagctttt taatttaatt 2150

attttaattc ttaatacttt aatttgtgca tttcataagc cccctgctct 2200

tggactgaat tttgtgcttt ttattgaaga attttattgt ttttatctta 2250

aaatcagttt ctattatcct tggggagacc atccctaaca aagtacaggt 2300

gggatctcct gtgagtcatt ggctgggttc tgattgctag atgtcacacc 2350

caccagcatc accaaagtga ctctgagata gaccggtccc ttctcagcgt 2400

tccagtcact tcaggaggaa tttagttatt gacttagtct atgacatctg 2450

gctacatgta ggtagagaag aaagacaatt ttaaaaagga aatcaggtct 2500

tttgcaactg tgcctccctc tgtctgtttt cacttgaatg ggtaaataac 2550

cagcagctag gttttgaatt cctaccttgt tattctaaac agatgtccac 2600

attgttaatt aaatctaaat tatgagcctt gctgagtgga tacggtactt 2650

acacctgaac caggattcct gggttctgtt gttgacattg cccttcagca 2700

cctgtttggc cagctgtata agataggact aatgactagg aagcctaccc 2750

caatgaatga tatactagat gaaatagtgt tcaaaacctg taggcactct 2800

ctggctaaaa acaaactctg aggccaccag cagatcatct ttaagctaag 2850

ttactatttt tcacctttt ttttagacgg agttttgctc tttgttgccc 2900

aggctggagt gcagtggcac gatctcggct cactgcaacc tccgcctccc 2950

aagttcaagc gattctcctg tctcagcctc ctgggtagct gggattacag 3000

gtgcccacca acatgcctgg ctaatttttg tacttttagt agagatgggg 3050

tttcaccatg ttggccaggc tggtcttcaa ctccagatct caggtgatct 3100

accctcctcg gcctcccaaa gtactgggat tacaggcctg agccaccgcg 3150

cccggcctat ttttcacttt aatttggcag ctgagaatgc ccaaaaagtg 3200

ccagaagcat cgtggcattt ccagaaccat ggattctgcc tttggacccc 3250

tctctattaa tattaaaact ctgggccttc agatgtcacc ctaatccact 3300

gccctaagac agaatttctg gacaagatgg gtaagggctt cattccttca 3350

acaagtcaag tcatacttgg cctctccctg agaatctgag caggagcctt 3400

ataacctgtg gtcattattt tttctttctg tacagaaata gaaaagcatt 3450

agaaataact tctaaccatc ctctgaaaaa acagaaaaaa tatcgaatcc 3500

ctctttcatg agaagtcttt tggataattg gaaaccttca tcactgaggt 3550

tggccagccc ctgccaagtg ttgtgtaggc aaagcacttg ttagtggctt 3600

cctatgaaat gttttagaga tctcttcacc atactggttt cttctctttg 3650

gttggtgtgg gtaaaagaaa acaaaacatt tcctataagc tgaaagctga 3700

ccagcattct cttcttggta acatctacta ctccaaccta gaaaatttgg 3750

attctagacc aaaaatcagg aaacatggct ccttataaat ctgtgcagct 3800

gccttatagt accatcaaag gaatttcagg tgggctgggc ggggccccga 3850

tcccagaatt atcaactcca cccatcatca tttggtcatg aagcatcctt 3900

tcattcttct tcttcttttt tttgggggggg gcggggcggg ggagggatct 3950

caaagtttta gtcttccaga atccaaatta aaggttgccc ctgatggggg 4000

ccaggttccg ccacagaaca tcttagatgt cagccttgac ctcacttagc 4050

agggattaca gaaatgagat acattttgaa ggagagttgt ctgttatgtt 4100

cactgtattc taagtgcctg ggataaagct gtctcatggg tgctccatat 4150

atattcatat atatttgttg agtgaattaa tgaattaaga gtggctggca 4200

gagtaggcag aaaaagacac tgcaaatggc ataaaaatta aagtcctagc 4250

tgagttctca atggtaaagg catcagatgt cttagcagtc aagctagaaa 4300

ttcatgacaa tgagtattac tatttgccta atgacaactc attgctctcc 4350

atgtaaatgt aatcaacaga tgaagagaat ataattgctc tgcttttcca 4400

ctaaaactcc atcttagtga attttaaatt atccagagat gtcaaactgc 4450

caaataaaaa tatttcagta gtctttgcat cagcttacct tgtaccagaa 4500

acatttccaa tttactatca aattatagta actgagcctg tgtgaagtat 4550

ctcatcattt tcgaaaggaa caccttgtgt gatgccagtg agcatttcta 4600

aaaagggtgt gaggtagagg taaaaataag gtgagagacc atttcagaat 4650

gcactgttgc tcaaaaaggt gatctggttc tttcttcaga gatttctacg 4700

gggatagaaa atcgggagtc tgccctcatt aatctgtgac tccacctctt 4750

gcatcaaatc aatatctatt tgttgagcac ttattgatta agaccttgca 4800

tatgtctgtc cattttgatt tgagatacaa ctttttgtgt gggttgaatg 4850

acaaatcact ccaaacaaaa ctgggcacag agaatcagct aggagaccag 4900

ttattcaggg tccatttctc ttggatgtaa aggagtcctg ggtaaaatgt 4950

ggctgtaacc taaaccaact agtccttgtg atttgtttct gccctctgtg 5000

tttcctgttg tcaaatgcta agtgtgtgtt ttgcagtcat gaactaaagc 5050

acaaaaagat gcatgagaca ttgtagtcat atgtctggtg tgacactttg 5100

gagcaaaaac cttgcagtgg taaataaaaa atttccaaca gggaaaaaaa 5150

aaaaaaaaaa aaaaaaaaaa 5170

<210> 63
<211> 5563
<212> DNA
<213> Homo sapiens

<400> 63
agcggcagct ccgctgaaaa ctgcattcag ccagtcctcc ggacttctgg 50

agcggggaca gggcgcaggg catcagcagc caccagcagg acctgggaaa 100

tagggattct tctgcctcca cttcaggttt tagcagcttg gtgctaaatt 150

gctgtctcaa aatgcagagg atctaatttg cagaggaaaa cagccaaaga 200

aggaagagga ggaaaaggaa aaaaaaaggg gtatattgtg gatgctctac 250

ttttcttgga aatgcaaaag attatgcata tttctgtcct cctttctcct 300

gttttatggg gactgatttt tggtgtctct tctaacagca tacagatagg 350

ggggctattt cctaggggcg ccgatcaaga atacagtgca tttcgagtag 400

ggatggttca gttttccact tcggagttca gactgacacc ccacatcgac 450

aatttggagg tggcaaacag cttcgcagtc actaatgctt tctgctccca 500

gttttcgaga ggagtctatg ctatttttgg attttatgac aagaagtctg 550

taaataccat cacatcattt tgcggaacac tccacgtctc cttcatcact 600

cccagcttcc caacagatgg cacacatcca tttgtcattc agatgagacc 650

cgacctcaaa ggagctctcc ttagcttgat tgaatactat caatgggaca 700

agtttgcata cctctatgac agtgacagag gcttatcaac actgcaagct 750

gtgctggatt ctgctgctga aaagaaatgg caagtgactg ctatcaatgt 800

gggaaacatt aacaatgaca agaaagatga gatgtaccga tcacttttc 850

aagatctgga gttaaaaaag gaacggcgtg taattctgga ctgtgaaagg 900

gataaagtaa acgacattgt agaccaggtt attaccattg gaaaacacgt 950

taaagggtac cactacatca ttgcaaatct ggaatttact gatggagacc 1000

tattaaaaat ccagtttgga ggtgcaaatg tctctggatt tcagatagtg 1050

gactatgatg attcgttggt atctaaattt atagaaagat ggtcaacact 1100

ggaagaaaaa gaataccctg gagctcacac aacaacaatt aagtatactt 1150

ctgctctgac ctatgatgcc gttcaagtga tgactgaagc cttccgcaac 1200

ctaaggaagc aaagaattga aatctcccga agggggaatg caggagactg 1250

tctggcaaac ccagcagtgc cctggggaca aggtgtagaa atagaaaggg 1300

ccctcaaaca ggttcaggtt gaaggtctct caggaaatat aaagtttgac 1350

cagaatggaa aaagaataaa ctatacaatt aacatcatgg agctcaaaac 1400

taatgggccc cggaagattg gctactggag tgaagtggac aaaatggttg 1450

ttacccttac tgagctccct tctggaaatg acacctctgg gcttgagaat 1500

aagactgttg ttgtcaccac aattttggaa tctccgtatg ttatgatgaa 1550

gaaaaatcat gaaatgcttg aaggcaatga gcgctatgag ggctactgtg 1600

ttgacctggc tgcagaaatc gccaaacatt gtgggttcaa gtacaagttg 1650

acaattgttg gtgatggcaa gtatggggcc agggatgcag acacgaaaat 1700

ttggaatggg atggttggag aacttgtata tgggaaagct gatattgcaa 1750

ttgctccatt aactattacc cttgtgagag aagaggtgat tgacttctca 1800

aagcccttca tgagcctcgg gatatctatc atgatcaaga agcctcagaa 1850

gtccaaacca ggagtgtttt cctttcttga tcctttagcc tatgagatct 1900

ggatgtgcat tgtttttgcc tacattgggg tcagtgtagt tttattcctg 1950

gtcagcagat ttagccccta cgagtggcac actgaggagt ttgaagatgg 2000

aagagaaaca caaagtagtg aatcaactaa tgaatttggg atttttaata 2050

gtctctggtt ttccttgggt gcctttatgc ggcaaggatg cgatatttcg 2100

ccaagatccc tctctgggcg cattgttgga ggtgtgtggt ggttctttac 2150

cctgatcata atctcctcct acacggctaa cttagctgcc ttcctgactg 2200

tagagaggat ggtgtctccc atcgaaagtg ctgaggatct ttctaagcaa 2250

acagaaattg cttatggaac attagactct ggctccacta aagagttttt 2300

caggagatct aaaattgcag tgtttgataa aatgtggacc tacatgcgga 2350

gtgcggagcc ctctgtgttt gtgaggacta cggccgaagg ggtggctaga 2400

gtgcggaagt ccaaagggaa atatgcctac ttgttggagt ccacgatgaa 2450

cgagtacatt gagcaaagga agccttgcga caccatgaaa gttggtggaa 2500

acctggattc caaaggctat ggcatcgcaa cacctaaagg atcctcatta 2550

ggaaccccag taaatcttgc agtattgaaa ctcagtgagc aaggcgtctt 2600

agacaagctg aaaaacaaat ggtggtacga taaaggtgaa tgtggagcca 2650

aggactctgg aagtaaggaa aagaccagtg ccctcagtct gagcaacgtt 2700

gctggagtat tctacatcct tgtcgggggc cttggtttgg caatgctggt 2750

ggctttgatt gagttctgtt acaagtcaag ggccgaggcg aaacgaatga 2800

aggtggcaaa gaatgcacag aatattaacc catcttcctc gcagaattca 2850

cagaattttg caacttataa ggaaggttac aacgtatatg gcatcgaaag 2900

tgttaaaatt taggggatga ccttgaatga tgccatgagg aacaaggcaa 2950

ggctgtcaat tacaggaagt actggagaaa atggacgtgt tatgactcca 3000

gaatttccca aagcagtgca tgctgtccct tacgtgagtc ctggcatggg 3050

aatgaatgtc agtgtgactg atctctcgtg attgataaga acctttgag 3100

tgccttacac aatggttttc ttgtgtgttt attgtcaaag tggtgagagg 3150

catccagtat cttgaagact tttctttcag ccaagaattc ttaaatatgt 3200

ggagttcatc ttgaattgta aggaatgatt aattaaaaca caacatcttt 3250

ttctactcga gttacagaca aagcgtggtg gacatgcaca gctaacatgg 3300

aagtactata atttacctga agtctttgta cagacaacaa acctgtttct 3350

gcagccacta ttgttagtct cttgattcat aatgacttaa gcacacttga 3400

catcaactgc atcaagatgt gacatgtttt ataaaaaaag gaaaaaaaac 3450

atttaaaact aaaaaatatt tttaggtatt ttcacaaaca aactggcttt 3500

taaataaatt tgcttccata ttggttgaat aagacaaaaa caattaaact 3550

gagtgggaag tgaataaaaa aaggctttag gtatcgattc catatttttc 3600

aaagccaaat atgtaaatgc taaggaaagt aaacaaagag gagattccaa 3650

tcttgtaatt taatattgtt attaaaactt taatgtatcc tattctttaa 3700

catttggtgt taatataaaa ttacttggca atgcttgaca tttgaaataa 3750

acatttttct attgttttat tgcaagtggt ccaattaatt ttgcttagct 3800

acagtttggt cataaatcaa gtgagtttaa agacactacc aagttgttag 3850

gtgcccagag aaaatttctc ccttttaaaa aggccaggtg atttttcaaa 3900

tgtaatcttg cccccaaagt aatatccgaa tatctttttg acatgtctaa 3950

atatatatat atataaagaa atatttgtta acacaaaagc atttgatcta 4000

tgtagataaa tgctaataga tttaaaaagc taatattaac aaataccaga 4050

atacgtgaag ttccattttt aaagtgtttg agcttacaga agagaaacat 4100

tcattttaaa tgaagtaaaa aatgccttga aagtaattct ttagatagtt 4150

gcccattgat taaattccaa aaactaaata tgtttttagc tttaaaatta 4200

taaaagctgt cataaacttt atatattatg aattttaaaa tatgtttgag 4250

tctcctgcaa tatagtttca tcccattgac atcaattaaa aataacccta 4300

atatattatt tttatattta ttcctcaggt ggaatggcta ttttaatatg 4350

cccagtgtgg ataaaatgtc acatttctgt aacttttgac taaagagcct 4400

atatttatct agttaatgaa tttaaaggat ctatctttcc cttcataaaa 4450

tacctcttat ttccattaaa gccccccaag tttaattaat ttaggatttt 4500

gaatgattat tgacatccaa tagttatttt taatatttgt attcttgtta 4550

tttctggaag aaagcctttg tgtagcactt ggtattttgc aaagtgcttt 4600

taaaacattc ttacttaccg tatttcatag aagggaagga aaaatgtaag 4650

gtttaacagt aagcacttgc attgaacatg gaggcatgtg gtatcatgat 4700

attcttcact aaatttagct gtccctaatc acagatccta aggtaatata 4750

atataatttt agtgcatttc tcctcatcag gaatgctgga ggtgcatttt 4800

aagttttaat aataagtgct agaatgacca aattgcagac taattgtttc 4850

catattgtac ttaaaatgag tttttaaaag tgaaaaagaa atgactatat 4900

acaatcaatg ctatttattg tacctctggg cctactcttc taaaaattgt 4950

agcttatcga tttttctctg tcaagcttga actaatgtaa ataattgaaa 5000

taatgtaaag ttatattttc atgtttttat agatacaaca tgacaagaat 5050

acataatgta agagtatttc aactatggat aatgttgatt ggataatgca 5100

catctcagtt acaagcagta ctcatagttt aatatccatg taacggtgca 5150

tcaatatatt gctatataaa tatgtctgtg tgcatataag tgaaaagtgg 5200

tcaaacaaga gtgatgacag ctgtctaaag gttttttttat tcattttata 5250

taaaaactgt tatggaaaga ccaaaatgtt tatgaactat tcttatgtaa 5300

atttacaatt gtcctttact gtacttttt gtttacagta tagtacctta 5350

ttttctgctg tgttaagtgg gtgtcaaact ccaagaagac atacactttc 5400

tataacttct attgaagata ttggaatttc caatttttca tgtgtactat 5450

gtcagaaaat gctttcgatt ttatttttaa atctaacatc ggatggcttt 5500

tccggagtgt tgtaaaaact tcaatcatac ataaaacatg ttcttacaaa 5550

aggcaaagat ctt 5563

<210> 64
<211> 1460
<212> DNA
<213> Homo sapiens

<400> 64
gccaacactg gccaaaaggg atagctgtcc aaggtctccc ccagcactga 50

ggagctcgcc tgctgccctc ttgcgcgcgg gaagcagcac caagttcacg 100

gccaacgcct tggcactagg gtccagaatg gctacaacag tccctgatgg 150

ttgccgcaat ggcctgaaat ccaagtacta cagactttgt gataaggctg 200

aagcttgggg catcgtccta gaaacggtgg ccacagccgg ggttgtgacc 250

tcggtggcct tcatgctcac tctcccgatc ctcgtctgca aggtgcagga 300

ctccaacagg cgaaaaatgc tgcctactca gtttctcttc ctcctgggtg 350

tgttgggcat ctttggcctc accttcgcct tcatcatcgg actggacggg 400

agcacagggc ccacacgctt cttcctcttt gggatcctct tttccatctg 450

cttctcctgc ctgctggctc atgctgtcag tctgaccaag ctcgtccggg 500

ggaggaagcc cctttccctg ttggtgattc tgggtctggc cgtgggcttc 550

agcctagtcc aggatgttat cgctattgaa tatattgtcc tgaccatgaa 600

taggaccaac gtcaatgtct tttctgagct ttccgctcct cgtcgcaatg 650

aagactttgt cctcctgctc acctacgtcc tcttcttgat ggcgctgacc 700

ttcctcatgt cctccttcac cttctgtggt tccttcacgg gctggaagag 750

acatggggcc cacatctacc tcacgatgct cctctccatt gccatctggg 800

tggcctggat caccctgctc atgcttcctg actttgaccg caggtgggat 850

gacaccatcc tcagctccgc cttggctgcc aatggctggg tgttcctgtt 900

ggcttatgtt agtcccgagt tttggctgct cacaaagcaa cgaaacccca 950

tggattatcc tgttgaggat gctttctgta aacctcaact cgtgaagaag 1000

agctatggtg tggagaacag agcctactct caagaggaaa tcactcaagg 1050

ttttgaagag acaggggaca cgctctatgc cccctattcc acacattttc 1100

agctgcagaa ccagcctccc caaaaggaat tctccatccc acgggcccac 1150

gcttggccga gcccttacaa agactatgaa gtaaagaaag agggcagcta 1200

actctgtcct gaagagtggg acaaatgcag ccgggcggca gatctagcgg 1250

gagctcaaag ggatgtgggc gaaatcttga gtcttctgag aaaactgtac 1300

aagacactac gggaacagtt tgcctccctc ccagcctcaa ccacaattct 1350

tccatgctgg ggctgatgtg ggctagtaag actccagttc ttagaggcgc 1400

tgtagtattt ttttttttt gtctcatcct ttagatactt cttttaagtg 1450

ggagtctcag 1460

<210> 65
<211> 1304
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 1144-1145
<223> unknown base

<400> 65
tcattccatt aaacatattt ctaaataata gtaagtggta ctaacaaaat 50

aaataataat ttaatagcct tagaaataaa tgactgtata cttatacagg 100

ttgaaaaaaa ctcggtagga ataagttacc tttttgttta ctaatgttgg 150

tttcaaaaat actcagattc attttagttg gctgacatct ggaagtagtt 200

aacaactaac cagtggactt caacaatcat ttgctcccag gcttccccca 250

tcatcaccct caccacatat cctgctaata tccaacaaca aacaaatatt 300

taatattgaa atagcccatt gcctgagaat gaacacaagc taaaatacat 350

gcaagggtac ttaatggaag ccaaaccatg ttctatacct aaggagaaaa 400

catggacatg tagaatgctt ttattcatgt attcaaaatc agaacaaatc 450

agtgtatatc actagaacat cagatggagg ataacacaag aagtgataca 500

gatcagggtt cacttctctt accctctctc tgttaggacc acattcctat 550

tttagccaaa tgtttctggt acgggccatc ttttcaccat aaatggcatt 600

atgtttcaaa tggctaaaag cttattatgg gtatgctcaa aggagtaaaa 650

cccaattcac agagatgtgg cttttctaaa gaaccaaatt aagggaatac 700

ttagtatcta caactaaagt attagtcatt ctgaggatta tttgttgctt 750

taagtattag ctccctacta catccataac accctactta ctaaatttaa 800

ttacacacaa cttttcaaga attcgtattt tatttgaagg gaggtacctg 850

tctactttat ctacaataaa aacaaagggt attggctttc tctaatccat 900

gcaaactaca aattccatcg ggagtcctac atcactaaca gtggtatgaa 950

caaaactaag aaagtacttc ctatccacat gtgaatgttt taaaaaagtt 1000

tttgccataa aacctaagtg taatttagca taccacagtg ctctgaagat 1050

gggtcattga cgatgtacca tttgtatata ggtaatacac atgtaaatca 1100

ctaattgtta attaaatata agaagtaggt ttttatcttt ttannaaaaa 1150

aacaaaaaag gtgactccct ttctcattct gttctgttgt attgcgtcca 1200

aaagttgtgt gtaatttttt taaggtccaa gaaatgtaaa gaaatttgtt 1250

ggaatacctg aatttctgta aaaaaataaa aataaaaata agattttgtt 1300

actt 1304

<210> 66
<211> 3268
<212> DNA
<213> Homo sapiens

<400> 66
gggaggctgt gccaggcgag ccggaggggt gctccgcgct cccccgccct 50

ccttccggga gcgaggatgc agactctgaa actggtgctg ctgggctgag 100

gcggaggcag gggagttgca gcgcgcgagg ctccgtgagt gtgtctcctg 150

cgcgctgaga ggcgggggga ggcggaggac caggaggagg aggaggagga 200

ggaggagggg gagaatgccc ggagccgccg ccgctgccgc cgccgccgcc 250

gccgcgatgc tcccggctca ggaggctgcc aagctgtacc acaccaacta 300

tgtgcggaac tcgcgggcca tcggcgtgct gtgggccatc ttcaccatct 350

gctttgccat cgtcaacgtg gtgtgcttca tccagcccta ctggataggc 400

gacggcgtgg acaccccgca agccggctat ttcgggctct tccactactg 450

catcggcaac ggcttctccc gggagctgac ctgcaggggc agcttcacgg 500

acttctccac gctgccctcg ggcgccttca aagccgcctc cttctttatc 550

ggcctctcca tgatgctcat cattgcctgc atcatttgct ttaccctctt 600

cttcttctgc aacacggcca ctgtgtacaa gatatgtgcc tggatgcagc 650

tcacctccgc tgcctgcctt gtgcttggct gtatgatttt ccctgatggc 700

tgggactcag atgaagtaaa acggatgtgt ggagaaaaga cagacaagta 750

cactcttggg gcttgctcag tccgctgggc atacatcctg gctattattg 800

gaattttgga tgccctgatc ctctcatttc tagcatttgt gcttggtaat 850

cgacaagaca gcttgatggc agaggaactg aaggcagaaa acaaagttct 900

gctaagccaa tattctctag aatgagcaca aaacaaatcg aataacagct 950

aaacaaatcg aataacagct aaacgaatcg aataacagct tttgtacatc 1000

aacatcaaga aggaatacgc ctgagagaga tcagagtata tagatgaata 1050

tgaacaagaa tggaacattc acttgtcaac gcactttcta aatctagatc 1100

agcagagatg ggagtgattt tctggaaaga gatgtgatca tggattaaac 1150

accagctcat tggaaactca ttggatgaga tcagaaaacg ttcatgaaaa 1200

atcatattca ggaaataagg aagaggaata taaatgctct agagttaaca 1250

tgtaaaatat atacgtactg aggtttgtaa actgtccttt ttaaatcaaa 1300

ctgaaaacaa aaagctttaa cctttcaaca gaatttttaa aaaggcagtt 1350

agttctaaat tattcctatc tcaatagcca agaggctgat caagcgtcat 1400

ttattgagga agcatcttag aaaatgcctc tgaatgtttt cataggagcc 1450

gtgacctttg gttcttcatc tctaccattc atttacttca ctgtgtaatt 1500

agttacaacc actcagttat taagagacgt aacgcttcaa acttttacc 1550

aagtctgtgt tctgtttaat ctgtccatac aagttattac tgagaaagtg 1600

tttatgccat atactattac tccatcaagc tgtatattac aggaagtaca 1650

tctttacatc ataggttccc aagcaacata gatttcccta tctttcagga 1700

aacagcatca aggaactctg aaaaatatag aaaaagttca ttttcacctt 1750

ggaagctcac gtgtaatatt ataggctact atcaaataaa cacttttttt 1800

ctaattctcc ctagtatatg cataggaatt taatatactt tataaataag 1850

tatctaaaat gtctcctact tttttcctat ttctttgcca tacatgttat 1900

cagaaatcca tgtcttctat ttcccttact gatgggcact cattttatt 1950

tttttaaaaa tcattccatt aaacatattt ctaaataata gtaagtggta 2000

ctaacaaaat aaataataat ttaatagcct tagaaataaa tgactgtata 2050

cttatacagg ttgaaaaaaa ctcggtagga ataagttacc ttttgttta 2100

ctaatgttgg tttcaaaaat actcagattc attttagttg gctgacatct 2150

ggaagtagtt aacaactaac cagtggactt caacaatcat ttgctcccag 2200

gcttccccca tcatcaccct caccacatat cctgctaata tccaacaaca 2250

aacaaatatt taatattgaa atagcccatt gcctgagaat gaacacaagc 2300

taaaatacat gcaagggtac ttaatggaag ccaaaccatg ttctatacct 2350

aaggagaaaa catggacatg tagaatgctt ttattcatgt attcaaaatc 2400

agaacaaatc agtgtatatc actagaacat cagatggagg ataacacaag 2450

aagtgataca gatcagggtt cacttctctt accctctctc tgttaggacc 2500

acattcctat tttagccaaa tgtttctggt acgggccatc ttttcaccat 2550

aaatggcatt atgtttcaaa tggctaaaag catattatgg gtatgctcaa 2600

aggagtaaaa cccaattcac agagatgtgg cttttctaaa gaaccaaatt 2650

aagggaatac ttagtatcta caactaaagt attagtcatt ctgaggatta 2700

tttgttgctt taagtattag ctccctacta catccataac accctactta 2750

ctaaatttaa ttacacacaa cttttcaaga attcgtattt tatttgaagg 2800

gaggtacctg tctactttat ctacaataaa aacaaagggt attggctttc 2850

tctaatccat gcaaactaca aattccatcg ggagtcctac atcactaaca 2900

gtggtatgaa caaaactaag aaagtacttc ctatccacat gtgaatgttt 2950

taaaaaagtt tttgccataa aacctaagtg taatttagca taccacagtg 3000

ctctgaagat gggtcattga cgatgtacca tttgtatata ggtaatacac 3050

atgtaaatca ctaattgtta attaaatata agaagtaggt ttttatcttt 3100

ttaaaaaaaa aacaaaaaag gtgactccct ttctcattct gttctgttgt 3150

attgcgtcca aaagttgtgt gtaatttttt taaggtccaa gaaatgtaaa 3200

gaaatttgtt ggaatacctg aatttctgta aaaaaataaa aataaaaata 3250

agattttgtt acttaaaa 3268

<210> 67
<211> 3268
<212> DNA
<213> Homo sapiens

<400> 67
gggaggctgt gccaggcgag ccggaggggt gctccgcgct cccccgccct 50

ccttccggga gcgaggatgc agactctgaa actggtgctg ctgggctgag 100

gcggaggcag gggagttgca gcgcgcgagg ctccgtgagt gtgtctcctg 150

cgcgctgaga ggcgggggga ggcggaggac caggaggagg aggaggagga 200

ggaggagggg gagaatgccc ggagccgccg ccgctgccgc cgccgccgcc 250

gccgcgatgc tcccggctca ggaggctgcc aagctgtacc acaccaacta 300

tgtgcggaac tcgcgggcca tcggcgtgct gtgggccatc ttcaccatct 350

gctttgccat cgtcaacgtg gtgtgcttca tccagcccta ctggataggc 400

gacggcgtgg acaccccgca agccggctat ttcgggctct tccactactg 450

catcggcaac ggcttctccc gggagctgac ctgcaggggc agcttcacgg 500

acttctccac gctgccctcg ggcgccttca aagccgcctc cttctttatc 550

ggcctctcca tgatgctcat cattgcctgc atcatttgct ttaccctctt 600

cttcttctgc aacacggcca ctgtgtacaa gatatgtgcc tggatgcagc 650

tcacctccgc tgcctgcctt gtgcttggct gtatgatttt ccctgatggc 700

tgggactcag atgaagtaaa acggatgtgt ggagaaaaga cagacaagta 750

cactcttggg gcttgctcag tccgctgggc atacatcctg gctattattg 800

gaattttgga tgccctgatc ctctcatttc tagcatttgt gcttggtaat 850

cgacaagaca gcttgatggc agaggaactg aaggcagaaa acaaagttct 900

gctaagccaa tattctctag aatgagcaca aaacaaatcg aataacagct 950

aaacaaatcg aataacagct aaacgaatcg aataacagct tttgtacatc 1000

aacatcaaga aggaatacgc ctgagagaga tcagagtata tagatgaata 1050

tgaacaagaa tggaacattc acttgtcaac gcactttcta aatctagatc 1100

agcagagatg ggagtgattt tctggaaaga gatgtgatca tggattaaac 1150

accagctcat tggaaactca ttggatgaga tcagaaaacg ttcatgaaaa 1200

atcatattca ggaaataagg aagaggaata taaatgctct agagttaaca 1250

tgtaaaatat atacgtactg aggtttgtaa actgtccttt ttaaatcaaa 1300

ctgaaaacaa aaagctttaa cctttcaaca gaatttttaa aaaggcagtt 1350

agttctaaat tattcctatc tcaatagcca agaggctgat caagcgtcat 1400

ttattgagga agcatcttag aaaatgcctc tgaatgtttt cataggagcc 1450

gtgacctttg gttcttcatc tctaccattc atttacttca ctgtgtaatt 1500

agttacaacc actcagttat taagagacgt aacgcttcaa acttttacc 1550

aagtctgtgt tctgtttaat ctgtccatac aagttattac tgagaaagtg 1600

tttatgccat atactattac tccatcaagc tgtatattac aggaagtaca 1650

tctttacatc ataggttccc aagcaacata gatttcccta tctttcagga 1700

aacagcatca aggaactctg aaaaatatag aaaaagttca ttttcacctt 1750

ggaagctcac gtgtaatatt ataggctact atcaaataaa cacttttttt 1800

ctaattctcc ctagtatatg cataggaatt taatatactt tataaataag 1850

tatctaaaat gtctcctact tttttcctat ttctttgcca tacatgttat 1900

cagaaatcca tgtcttctat ttcccttact gatgggcact cattttatt 1950

ttttaaaaa tcattccatt aaacatattt ctaaataata gtaagtggta 2000

ctaacaaaat aaataataat ttaatagcct tagaaataaa tgactgtata 2050

cttatacagg ttgaaaaaaa ctcggtagga ataagttacc tttttgttta 2100

ctaatgttgg tttcaaaaat actcagattc attttagttg gctgacatct 2150

ggaagtagtt aacaactaac cagtggactt caacaatcat ttgctcccag 2200

gcttccccca tcatcaccct caccacatat cctgctaata tccaacaaca 2250

aacaaatatt taatattgaa atagcccatt gcctgagaat gaacacaagc 2300

taaaatacat gcaagggtac ttaatggaag ccaaaccatg ttctatacct 2350

aaggagaaaa catggacatg tagaatgctt ttattcatgt attcaaaatc 2400

agaacaaatc agtgtatatc actagaacat cagatggagg ataacacaag 2450

aagtgataca gatcagggtt cacttctctt accctctctc tgttaggacc 2500

acattcctat tttagccaaa tgtttctggt acgggccatc ttttcaccat 2550

aaatggcatt atgtttcaaa tggctaaaag catattatgg gtatgctcaa 2600

aggagtaaaa cccaattcac agagatgtgg cttttctaaa gaaccaaatt 2650

aagggaatac ttagtatcta caactaaagt attagtcatt ctgaggatta 2700

tttgttgctt taagtattag ctccctacta catccataac accctactta 2750

ctaaatttaa ttacacacaa cttttcaaga attcgtattt tatttgaagg 2800

gaggtacctg tctactttat ctacaataaa aacaaagggt attggctttc 2850

tctaatccat gcaaactaca aattccatcg ggagtcctac atcactaaca 2900

gtggtatgaa caaaactaag aaagtacttc ctatccacat gtgaatgttt 2950

taaaaaagtt tttgccataa aacctaagtg taatttagca taccacagtg 3000

ctctgaagat gggtcattga cgatgtacca tttgtatata ggtaatacac 3050

atgtaaatca ctaattgtta attaaatata agaagtaggt ttttatcttt 3100

ttaaaaaaaa aacaaaaaag gtgactccct ttctcattct gttctgttgt 3150

attgcgtcca aaagttgtgt gtaatttttt taaggtccaa gaaatgtaaa 3200

gaaatttgtt ggaatacctg aatttctgta aaaaaataaa aataaaaata 3250

agattttgtt acttaaaa 3268

<210> 68
<211> 3268
<212> DNA
<213> Homo sapiens

<400> 68

gggaggctgt gccaggcgag ccggaggggt gctccgcgct cccccgccct 50

ccttccggga gcgaggatgc agactctgaa actggtgctg ctgggctgag 100

gcggaggcag gggagttgca gcgcgcgagg ctccgtgagt gtgtctcctg 150

cgcgctgaga ggcgggggga ggcggaggac caggaggagg aggaggagga 200

ggaggagggg gagaatgccc ggagccgccg ccgctgccgc cgccgccgcc 250

gccgcgatgc tcccggctca ggaggctgcc aagctgtacc acaccaacta 300

tgtgcggaac tcgcgggcca tcggcgtgct gtgggccatc ttcaccatct 350

gctttgccat cgtcaacgtg gtgtgcttca tccagcccta ctggataggc 400

gacggcgtgg acaccccgca agccggctat ttcgggctct tccactactg 450

catcggcaac ggcttctccc gggagctgac ctgcaggggc agcttcacgg 500

acttctccac gctgccctcg ggcgccttca aagccgcctc cttctttatc 550

ggcctctcca tgatgctcat cattgcctgc atcatttgct ttaccctctt 600

cttcttctgc aacacggcca ctgtgtacaa gatatgtgcc tggatgcagc 650

tcacctccgc tgcctgcctt gtgcttggct gtatgatttt ccctgatggc 700

tgggactcag atgaagtaaa acggatgtgt ggagaaaaga cagacaagta 750

cactcttggg gcttgctcag tccgctgggc atacatcctg gctattattg 800

gaattttgga tgccctgatc ctctcatttc tagcatttgt gcttggtaat 850

cgacaagaca gcttgatggc agaggaactg aaggcagaaa acaaagttct 900

gctaagccaa tattctctag aatgagcaca aaacaaatcg aataacagct 950

aaacaaatcg aataacagct aaacgaatcg aataacagct tttgtacatc 1000

aacatcaaga aggaatacgc ctgagagaga tcagagtata tagatgaata 1050

tgaacaagaa tggaacattc acttgtcaac gcactttcta aatctagatc 1100

agcagagatg ggagtgattt tctggaaaga gatgtgatca tggattaaac 1150

accagctcat tggaaactca ttggatgaga tcagaaaacg ttcatgaaaa 1200

atcatattca ggaaataagg aagaggaata taaatgctct agagttaaca 1250

tgtaaaatat atacgtactg aggtttgtaa actgtccttt ttaaatcaaa 1300

ctgaaaacaa aaagctttaa cctttcaaca gaatttttaa aaaggcagtt 1350

agttctaaat tattcctatc tcaatagcca agaggctgat caagcgtcat 1400

ttattgagga agcatcttag aaaatgcctc tgaatgtttt cataggagcc 1450

gtgacctttg gttcttcatc tctaccattc atttacttca ctgtgtaatt 1500

agttacaacc actcagttat taagagacgt aacgcttcaa acttttacc 1550

aagtctgtgt tctgtttaat ctgtccatac aagttattac tgagaaagtg 1600

tttatgccat atactattac tccatcaagc tgtatattac aggaagtaca 1650

tctttacatc ataggttccc aagcaacata gatttcccta tctttcagga 1700

aacagcatca aggaactctg aaaaatatag aaaaagttca ttttcacctt 1750

ggaagctcac gtgtaatatt ataggctact atcaaataaa cactttttt 1800

ctaattctcc ctagtatatg cataggaatt taatatactt tataaataag 1850

tatctaaaat gtctcctact tttttcctat ttctttgcca tacatgttat 1900

cagaaatcca tgtcttctat ttcccttact gatgggcact cattttatt 1950

tttttaaaaa tcattccatt aaacatattt ctaaataata gtaagtggta 2000

ctaacaaaat aaataataat ttaatagcct tagaaataaa tgactgtata 2050

cttatacagg ttgaaaaaaa ctcggtagga ataagttacc tttttgttta 2100

ctaatgttgg tttcaaaaat actcagattc attttagttg gctgacatct 2150

ggaagtagtt aacaactaac cagtggactt caacaatcat ttgctcccag 2200

gcttccccca tcatcaccct caccacatat cctgctaata tccaacaaca 2250

aacaaatatt taatattgaa atagcccatt gcctgagaat gaacacaagc 2300

taaaatacat gcaagggtac ttaatggaag ccaaaccatg ttctatacct 2350

aaggagaaaa catggacatg tagaatgctt ttattcatgt attcaaaatc 2400

agaacaaatc agtgtatatc actagaacat cagatggagg ataacacaag 2450

aagtgataca gatcagggtt cacttctctt accctctctc tgttaggacc 2500

acattcctat tttagccaaa tgtttctggt acgggccatc ttttcaccat 2550

aaatggcatt atgtttcaaa tggctaaaag catattatgg gtatgctcaa 2600

aggagtaaaa cccaattcac agagatgtgg cttttctaaa gaaccaaatt 2650

aagggaatac ttagtatcta caactaaagt attagtcatt ctgaggatta 2700

tttgttgctt taagtattag ctccctacta catccataac accctactta 2750

ctaaatttaa ttacacacaa cttttcaaga attcgtattt tatttgaagg 2800

gaggtacctg tctactttat ctacaataaa aacaaagggt attggctttc 2850

tctaatccat gcaaactaca aattccatcg ggagtcctac atcactaaca 2900

gtggtatgaa caaaactaag aaagtacttc ctatccacat gtgaatgttt 2950

taaaaaagtt tttgccataa aacctaagtg taatttagca taccacagtg 3000

ctctgaagat gggtcattga cgatgtacca tttgtatata ggtaatacac 3050

atgtaaatca ctaattgtta attaaatata agaagtaggt ttttatcttt 3100

ttaaaaaaaa aacaaaaaag gtgactccct ttctcattct gttctgttgt 3150

attgcgtcca aaagttgtgt gtaatttttt taaggtccaa gaaatgtaaa 3200

gaaatttgtt ggaatacctg aatttctgta aaaaaataaa aataaaaata 3250

agattttgtt acttaaaa 3268

<210> 69
<211> 1755
<212> DNA
<213> Homo sapiens

<400> 69
caggcctctg aggctccctt gccgagggcc ccgagctgca gggacagtga 50

gcagtgagtc ccttgggcat cccgctcctg ggcaggtcac caataggtcc 100

ccgcagttcc caatggaact gttccagtcc tccccgaggc ctccacttca 150

acctgtctgt gtctgcccag gcctggagtt gtgtgaccct ccccaccgcc 200

tggccttctc catgggggct ggccttttct cggtggtggg caccctgctg 250

ctgcccggcc tggctgcgct tgtgcaggac tggcgtcttc tgcaggggct 300

gggtgccctg atgagtggac tcttgctgct cttttggggg ttcccggccc 350

tgttccccga gtctccctgc tggctgctgg ccacaggtca ggtagctcga 400

gccaggaaga tcctgtggcg ctttgcagaa gccagtggcg tgggccccgg 450

ggacagttcc ttggaggaga actccctggc tacagagctg accatgctgt 500

ctgcacggag cccccagccc cggtaccact ccccactggg gcttctgcgt 550

acccgagtca cctggagaaa cgggcttatc ttgggcttca gctcgctggt 600

tggtggaggc atcagagcta gcttccgccg cagcctggca cctcaggtgc 650

cgaccttcta cctgccctac ttcctggagg ccggcctgga ggcggcagcc 700

ttggtcttcc tgctcctgac ggcagattgc tgtggacgcc gccccgtgct 750

gctgctgggc accatggtca caggcctggc atccctgctg ctcctcgctg 800

gggcccagta tctgccaggc tggactgtgc tgttcctctc tgtcctgggg 850

ctcctggcct cccgggctgt gtccgcactc agcagcctct tcgcggccga 900

ggtcttcccc acggtgatca gggggggccgg gctgggcctg gtgctggggg 950

ccgggttcct gggccaggca gccggccccc tggacaccct gcacggccgg 1000

cagggcttct tcctgcaaca agtcgtcttc gcctcccttg ctgtccttgc 1050

cctgctgtgt gtcctgctgc tgcctgagag ccgaagccgg gggctgcccc 1100

agtcactgca ggacgccgac cgcctgcgcc gctccccact cctgcggggc 1150

cgcccccgcc aggaccacct gcctctgctg ccgccctcca actcctactg 1200

ggccggccac acccccgagc agcactagtc ctgcctggtg gccctgggag 1250

ccaggatggg accaaagtca aggcctgggg catggctgag taccccagac 1300

gtctggtcca gggcagacac attcctctca gaagcccgtg tctcagtgca 1350

ggtggagccg tggggacagc gtgaaggtgt ctccagccag gccccaggca 1400

ctgggaggcc ctgggtctcc ccccagccac acccagtagg tgtggaggat 1450

aaaggcttct gtggaacttg gcgtcctccc tggttcttcc ctcaggccac 1500

acctggggcc tggaagggtc tcacctgtgc ctgctcccca cctctgggct 1550

caccagcaag gagccagggt gtcagcggct gctccatagc ccctgtccag 1600

gtagtggcag cccccggccg cctcctgccc cagccaagca gggtggtatg 1650

ggggcctggc cctcaggtcc ctgccaggcc aggcctcact ggccccatcc 1700

ccctcatggt cagtgtggca gtgggcgccc gtcctgcatg aaagagacta 1750

gcacg 1755

<210> 70
<211> 4021
<212> DNA
<213> Homo sapiens

<400> 70
agcggggcgg gtccccgaag cccagtacca gacctcctgt gcgaagagta 50

acctgctctt cggactgata actggatgga atggccttta gaaagggagc 100

atcttcagaa tctgcagctg ggatcagatt gttcttttgt tgaatcaact 150

cagctctgcc gagccagcat cttgctaagg gttacaaaat aaatgccatt 200

tgaacagtgc catctgtcat ggaaaaacgt gagacgtttg tacaatgccg 250

tgtctaagga gctggttgga gagtttttgc aatttgttca acttgataaa 300

gaggcctctg atcctttcag cctaaatgaa ttactagatg aattatcaag 350

gaaacagaaa gaagaattat ggcaaaggct gaagaattta ttgacagatg 400

tgttgttaga aagcccagtg gatgggtggc aggtagtgga agcccagggt 450

gaagacaata tggaaaccga acatggctca aaaatgagaa aaagcataga 500

aataatttat gcaattacat ctgtgattct tgcttctgtg tctgtaataa 550

atgaaagtga gaactacgaa gccctactgg aatgtgttat tatattaaat 600

ggtattttat atgcattacc tgagtctgaa cgaaaactac agagttctat 650

tcaggatttg tgtgttacct ggtgggagaa aggcctgcct gccaaggaag 700

acacaggaaa gacagccttt gtcatgttac taaggaggag tctggagact 750

aagacaggtg cagacgtatg tcggctttgg cgtatccatc aagctttata 800

ttgctttgat tatgatttgg aggaaagtgg agaaattaaa gatatgttac 850

ttgagtgctt cataaatatt aattatatca agaaagaaga gggaagaaga 900

tttcttagtt gtctcttcaa ctggaatatc aacttcatca aaatgatcca 950

cgggaccatt aaaaaccagt tacagggatt acaaaagtct ttgatggtat 1000

acattgcaga aatttatttc agagcttgga aaaaggcttc agggaaaata 1050

ctggaggcga ttgaaaatga ttgcatccag gacttcatgt tccacgggat 1100

acaccttccg aggaggtctc cagtgcattc caaagtgcgg gaggttctga 1150

gttactttca ccatcaaaag aaagttcggc agggagtgga agagatgctt 1200

tatagattat ataagcccat cctttggaga ggattaaagg ccagaaactc 1250

tgaagttcga tcaaatgctg cattgttgtt tgttgaagca tttcctatta 1300

gggatccaaa ccttcatgct attgaaatgg atagtraaat ccagaaacag 1350

tttgaagagc tctatagcct tttagaagat ccttacccga tggtccgttc 1400

cacagggatc cttggtgttt gtaaaataac ttctaagtac tgggaaatga 1450

tgcccccgac cattcttatt gacctcctga agaaggtgac tggggaactg 1500

gcatttgaca cgagctcagc tgatgttcgt tgttctgtct ttaagtgtct 1550

gccaatgatt ttggacaaca aactgagcca cccattgtta gagcagctcc 1600

ttccagctct cagatacagt ctccacgaca attcggagaa agtgagggta 1650

gcttttgtgg acatgctgtt gaagatcaaa gctgtgaggg ctgctaagtt 1700

ttggaaaata tgtcccatgg agcacattct ggttcgtctg gaaactgatt 1750

ctcgacctgt gtctcggcgc ctggtgagcc tcatctttaa ttctttcctg 1800

cctgtgaatc agccggagga ggtctggtgc gagcgctgtg tcaccctggt 1850

gcagatgaac cacgccgctg ccaggaggtt ctatcagtac gcccacgaac 1900

acaccgcctg caccaacata gcaaagctga ttcacgttat tcgtcattgc 1950

ttaaatgcct gtatccagag ggcagtgaga gagcctccag aggacgagga 2000

ggaagaggac ggaagggaga aggagaatgt gactgttctg gacaaaacac 2050

tgtcagtaaa cgatgttgca tgcatggcag gtttactaga aatcattgtg 2100

attctctgga aaagtattga cagaagtatg gaaaataata aagaggccaa 2150

actttacacg attaacaagt ttgcctctgt gcttccagag tatctgaaag 2200

tatttaagga tgatcgctgc aagatccctt tattcatgct aatgtccttt 2250

atgccggcct ctgctgtccc cccattcagc tgtggtgtga tttccacgct 2300

gagaagccgg gaggagggcg ctgtggacaa gagctactgc actttgttgg 2350

attgcctctg ctcctggggg caggtggggc acattctgga gcttgttgac 2400

aactggctgc ccacagagca tgcccaggcc aagagcaaca cagcttctaa 2450

aggtagggtg cagatccatg acacacgccc agtcaaacct gaattggcat 2500

tggtctacat tgagtatctg ctgactcatc caaagaaccg cgagtgcttg 2550

ctctctgctc ctcggaagaa acttaaccat cttttgaaag cccttgaaac 2600

gtcaaaggca gatctggagt cacttctgca gacaccgggt gggaagcctc 2650

gtggcttcag tgaagcagct gccccgcgag cctttggtct ccactgtcgc 2700

ctgagcatcc atcttcagca caagttctgc tcagaaggaa aggtgtattt 2750

gtccatgttg gaagacactg gcttttggtt agaaagcaaa attttatctt 2800

ttattcaaga tcaagaagaa gactacctga agcttcatag ggtcatttat 2850

cagcaaatta tccagaccta cctgactgtg tgtaaagatg ttgttatggt 2900

aggccttggt gaccatcagt ttcagatgca actcttacag cggagtcttg 2950

gaatcatgca aacagtgaag ggatttttt atgtttcatt acttcttgac 3000

attctgaaag agataactgg aagttccttg attcagaaaa cagattcaga 3050

tgaagaagtt gcaatgctgt tggacacagt ccagaaagta tttcagaaaa 3100

tgttggaatg tattgcacgg agcttcagga agcagccgga agaaggcctg 3150

cggctgcttt attctgttca gaggcctctt catgagttca ttactgctgt 3200

tcagtctcgg cacacagaca cccctgtgca ccggggtgta ctttctactc 3250

tgatcgctgg gcctgtggtt gagataagtc accagctacg gaaggtttct 3300

gacgtagaag agcttacccc tccagagcat ctttctgatc ttccaccatt 3350

ttcaaggtgt ttaataggaa taataataaa gtcttcgaat gtggtcaggt 3400

catttttgga tgaattaaag gcatgtgtgg cttctaatga tattgaaggc 3450

attgtgtgcc tcacggctgc tgtgcatatt atcctggtta ttaatgcagg 3500

taaacataaa agctcaaaag tgagggaggt tgcagccact gttcacagaa 3550

aactaaagac attcatggaa attactttgg aagaggatag cattgaaaga 3600

tttctctatg aatcatcatc aagaactctg ggagaacttt tgaattcata 3650

accaagccaa catctccaga catgtaaaaa tagggaaaag tgattcaaat 3700

tgaaatgcct gtgtattttc ctattgtttt taatgttaat aacccatata 3750

atagggaaag ggtgggattt ttttgtggga atgtgggaag gtgggggtta 3800

tggaggagat aactcaaaac ttcttcaatt ttgcctagtg cctgcgtaaa 3850

taatatattt aatataaagg actccaggta tgaatggtgt agaaatccat 3900

gattccaaga aaaaacactt ttctagcaaa cctggttgtt tttaaaatga 3950

cttttatata tgtaatattg cttggaaact atgagtaata aagcaatgac 4000

aacaamaaaa aaaaaaaaaa a 4021

<210> 71
<211> 3580
<212> DNA
<213> Homo sapiens

<400> 71
gcccttcaga atctgcagct gggatcagat tgttcttttg ttgaatcaac 50

tcagctctgc cgagccagca tcttgctaag ggttacaaaa taaatgccat 100

ttgaacagtg ccatctgtca tggaaaaacg tgagacgttt gtacaagccg 150

tgtctaagga gctggttgga gagtttttgc aatttgttca acttgataaa 200

gaggcctctg atcctttcag cctaaatgaa ttactagatg aattatcaag 250

gaaacagaaa gaagaattat ggcaaaggct gaagaattta ttgacagatg 300

tgttgttaga aagcccagtg gatgggtggc aggtagtgga agcccagggt 350

gaagacaata tggaaaccga acatggctca aaaatgagaa aaagcataga 400

aataatttat gcaattacat ctgtgattct tgcttctgtg tctgtaataa 450

atgaaagtga gaactacgaa gccctactgg aatgtgttat tatattaaat 500

ggtattttat atgcattacc tgagtctgaa cgaaaactac agagttctat 550

tcaggatttg tgtgttacct ggtgggagaa aggcctgcct gccaaggaag 600

acacaggaaa gacagccttt gtcatgttac taaggaggag tctggagact 650

aagacaggtg cagacgtatg tcggctttgg cgtatccatc aagctttata 700

ttgctttgat tatgatttgg aggaaagtgg agaaattaaa gatatgttac 750

ttgagtgctt cataaatatt aattatatca agaaagaaga gggaagaaga 800

tttcttagtt gtctcttcaa ctggaatatc aacttcatca aaatgatcca 850

cgggaccatt aaaaaccagt tacagggatt acaaaagtct ttgatggtat 900

acattgcaga aatttatttc agagcttgga aaaaggcttc agggaaaata 950

ctggaggcga ttgaaaatga ttgcatccag gacttcatgt tccacgggat 1000

acaccttccg aggaggtctc cagtgcattc caaagtgcgg gaggttctga 1050

gttactttca ccatcaaaag aaagttcggc agggagtgga agagatgctt 1100

tatagattat ataagcccat cctttggaga ggattaaagg ccagaaactc 1150

tgaagttcga tcaaatgctg cattgttgtt tgttgaagca tttcctatta 1200

gggatccaaa ccttcatgct attgaaatgg atagtgaaat ccagaaacag 1250

tttgaagagc tctatagcct tttagaagat ccttacccga tggtccgttc 1300

cacagggatc cttggtgttt gtaaaataac ttctaagtac tgggaaatga 1350

tgcccccgac cattcttatt gacctcctga agaaggtgac tggggaactg 1400

gcatttgaca cgagctcagc tgatgttcgt tgttctgtct ttaagtgtct 1450

gccaatgatt ttggacaaca aactgagcca cccattgtta gagcagctcc 1500

ttccagctct cagatacagt ctccacgaca attcggagaa agtgagggta 1550

gcttttgtgg acatgctgtt gaagatcaaa gctgtgaggg ctgctaagtt 1600

ttggaaaata tgtcccatgg agcacattct ggttcgtctg gaaactgatt 1650

ctcgacctgt gtctcggcgc ctggtgagcc tcatctttaa ttctttcctg 1700

cctgtgaatc agccggagga ggtctggtgc gagcgctgtg tcaccctggt 1750

gcagatgaac cacgccgctg ccaggaggtt ctatcagtac gcccacgaac 1800

acaccgcctg caccaacata gcaaagctga ttcacgttat tcgtcattgc 1850

ttaaatgcct gtatccagag ggcagtgaga gagcctccag aggacgagga 1900

ggaagaggac ggaagggaga aggagaatgt gactgttctg gacaaaacac 1950

tgtcagtaaa cgatgttgca tgcatggcag gtttactaga aatcattgtg 2000

attctctgga aaagtattga cagaagtatg gaaaataata aagaggccaa 2050

actttacacg attaacaagt ttgcctctgt gcttccagag tatctgaaag 2100

tatttaagga tgatcgctgc aagatccctt tattcatgct aatgtccttt 2150

atgccggcct ctgctgtccc cccattcagc tgtggtgtga tttccacgct 2200

gagaagccgg gaggagggcg ctgtggacaa gagctactgc actttgttgg 2250

attgcctctg ctcctggggg caggtggggc acattctgga gcttgttgac 2300

aactggctgc ccacagagca tgcccaggcc aagagcaaca cagcttctaa 2350

aggtagggtg cagatccatg acacacgccc agtcaaacct gaattggcat 2400

tggtctacat tgagtatctg ctgactcatc caaagaaccg cgagtgcttg 2450

ctctctgctc ctcggaagaa acttaaccat cttttgaaag cccttgaaac 2500

gtcaaaggca gatctggagt cacttctgca gacaccgggt gggaagcctc 2550

gtggcttcag tgaagcagct gccccgcgag cctttggtct ccactgtcgc 2600

ctgagcatcc atcttcagca caagttctgc tcagaaggaa aggtgtattt 2650

gtccatgttg gaagacactg gcttttggtt agaaagcaaa attttatctt 2700

ttattcaaga tcaagaagaa gactacctga agcttcatag ggtcatttat 2750

cagcaaatta tccagaccta cctgactgtg tgtaaagatg ttgttatggt 2800

aggccttggt gaccatcagt ttcagatgca actcttacag cggagtcttg 2850

gaatcatgca aacagtgaag ggatttttt atgtttcatt acttcttgac 2900

attctgaaag agataactgg aagttccttg attcagaaaa cagattcaga 2950

tgaagaagtt gcaatgctgt tggacacagt ccagaaagta tttcagaaaa 3000

tgttggaatg tattgcacgg agcttcagga agcagccgga agaaggcctg 3050

cggctgcttt attctgttca gaggcctctt catgagttca ttactgctgt 3100

tcagtctcgg cacacagaca cccctgtgca ccggggtgta ctttctactc 3150

tgatcgctgg gcctgtggtt gagataagtc accagctacg gaaggtttct 3200

gacgtagaag agcttacccc tccagagcat ctttctgatc ttccaccatt 3250

ttcaaggtgt ttaataggaa taataataaa gtcttcgaat gtggtcaggt 3300

catttttgga tgaattaaag gcatgtgtgg cttctaatga tattgaaggc 3350

attgtgtgcc tcacggctgc tgtgcatatt atcctggtta ttaatgcagg 3400

taaacataaa agctcaaaag tgagggaggt tgcagccact gttcacagaa 3450

aactaaagac attcatggaa attactttgg aagaggatag cattgaaaga 3500

tttctctatg aatcatcatc aagaactctg ggagaacttt tgaattcata 3550

accaagccaa catctccaga catgaagggc 3580

<210> 72
<211> 1191
<212> DNA
<213> Homo sapiens

<400> 72
tccaaccaga ttttcaccgc tatgcctcct ttcactctgg gagtcttcca 50

gaggtcttgc actcgggaga gcatgctcag gtttccccag ctctacaaaa 100

tcacccagaa tgccaaagac ttcaacgtga ccagtggtca tgccactgcc 150

tgttgatttg ttgaaaatat tgtttacacg tatgttcttg ttactgattg 200

tcagaaagct ggttttgaga ctgcagcttg gactaaattc agtcatctgg 250

ctgtctgggg aagcatgctg accagtctgg tgttctttgg catctactca 300

gccatctggt ccaccattct cattgcccca aatatgagag gacagaagaa 350

tggtaccggt actgccaatg gagatggagg aaggagacag aaagaaacag 400

agcccagacc ctagggacca ccagcatttg cagaatggat aaacagcctt 450

cttcctaaca aaggaagcac agcaactgtg atcctgagct gtgcacactt 500

ctggttggga ttatttctgg tttctacttc ctgtttgaag atgtggcatg 550

gagagtgaac aagctgctgc ccaccacctg gcatcacagc cccagaactc 600

agctatttcc atgggaccac agcatctcat ctctgggctg agccagaaag 650

acccctactg aagtccagag gcacttttct gaaaggctct gctttgacct 700

gaagtatttt atctatcctc agtctcagga cactgttgat ggaattaagg 750

ccaagcacat ctgcaaaaaa gacattgctg gaggaggtgc aaagagctgg 800

aaaccaagtc tccagtcctg ggaaaagcag tggtatggaa aagcaatgga 850

aagagcattt tgaaaatgcc attccactgt tttctggcct ttatgatttc 900

tgctgagaaa tccactgtta gtctgatggg gtctccttca tagcaccaat 950

gacctgaaga gccttgttga aggaagactc catctgatga ctcagagcaa 1000

gtatttttta gtgtgttatt gttattagca gaaagagggc cataaaatac 1050

atggggcaag ctgaatatat cttaggcaaa agaagaaaat attcaaattc 1100

ttatgttatt ttatctaatt attttatctc tttttgtgtg tgacttataa 1150

tgtgtgtatt gtattaataa aagtatataa acatgtagtt t 1191

<210> 73
<211> 2752
<212> DNA
<213> Homo sapiens

<400> 73
cattcccacc cttccttctt taataagcag gagcgaaaaa gacaaattcc 50

aaagaggatt gttcagttca agggaatgaa gaattcagaa taattttggt 100

aaatggattc caatatcggg aataagaata agctgaacag ttgacctgct 150

ttgaagaaac atactgtcca tttgtctaaa ataatctata acaaccaaac 200

caatcaaaat gaattcaaca ttattttccc aggttgaaaa tcattcagtc 250

cactctaatt tctcagagaa gaatgcccag cttctggctt ttgaaaatga 300

tgattgtcat ctgcccttgg ccatgatatt taccttagct cttgcttatg 350

gagctgtgat cattcttggt gtctctggaa acctggcctt gatcataatc 400

atcttgaaac aaaaggagat gagaaatgtt accaacatcc tgattgtgaa 450

cctttccttc tcagacttgc ttgttgccat catgtgtctc ccctttacat 500

ttgtctacac attaatggac cactgggtct ttggtgaggc gatgtgtaag 550

ttgaatcctt ttgtgcaatg tgtttcaatc actgtgtcca ttttctctct 600

ggttctcatt gctgtggaac gacatcagct gataatcaac cctcgagggt 650

ggagaccaaa taatagacat gcttatgtag gtattgctgt gatttgggtc 700

cttgctgtgg cttcttcttt gcctttcctg atctaccaag taatgactga 750

tgagccgttc caaaatgtaa cacttgatgc gtacaaagac aaatacgtgt 800

gctttgatca atttccatcg gactctcata ggttgtctta taccactctc 850

ctcttggtgc tgcagtattt tggtccactt tgttttatat ttatttgcta 900

cttcaagata tatatacgcc taaaaaggag aaacaacatg atggacaaga 950

tgagagacaa taagtacagg tccagtgaaa ccaaaagaat caatatcatg 1000

ctgctctcca ttgtggtagc atttgcagtc tgctggctcc ctcttaccat 1050

ctttaacact gtgtttgatt ggaatcatca gatcattgct acctgcaacc 1100

acaatctgtt attcctgctc tgccacctca cagcaatgat atccacttgt 1150

gtcaacccca tattttatgg gttcctgaac aaaaacttcc agagagactt 1200

gcagttcttc ttcaactttt gtgatttccg gtctcgggat gatgattatg 1250

aaacaatagc catgtccacg atgcacacag atgtttccaa aacttctttg 1300

aagcaagcaa gcccagtcgc atttaaaaaa atcaacaaca atgatgataa 1350

tgaaaaaatc tgaaactact tatagcctat ggtcccggat gacatctgtt 1400

taaaaacaag cacaacctgc aacatacttt gattacctgt tctcccaagg 1450

aatggggttg aaatcatttg aaaatgacta agattttctt gtcttgcttt 1500

ttactgcttt tgttgtagtt gtcataatta catttggaac aaaaggtgtg 1550

ggctttgggg tcttctggaa atagttttga ccagacatct ttgaagtgct 1600

ttttgtgaat ttatgcatat aatataaaga cttttatact gtacttattg 1650

gaatgaaatt tctttaaagt attacgatgc gctgacttca gaagtacctg 1700

ccatccaata cggtcattag attgggtcat cttgattaga ttagattaga 1750

ttagattgtc aacagattgg gccatcctta ctttatgata ggcatcattt 1800

tagtgtgtta caatagtaac agtatgcaaa agcagcattc aggagccgaa 1850

agatagtctt gaagtcattc agaagtggtt tgaggtttct gtttttggt 1900

ggtttttgtt tgttttttt ttttttcacc ttaagggagg ctttcatttc 1950

ctcccgactg attgtcactt aaatcaaaat ttaaaaatga ataaaaagac 2000

atacttctca gctgcaaata ttatggagaa ttgggcaccc acaggaatga 2050

agagagaaag cagctcccca acttcaaaac cattttggta cctgacaaca 2100

agagcatttt agagtaatta atttaataaa gtaaattagt attgctgcaa 2150

atagctaaat tatatttatt tgaattgatg gtcaagagat tttccatttt 2200

ttttacagac tgttcagtgt ttgtcaagct tctggtctaa tatgtactcg 2250

aaagactttc cgcttacaat ttgtagaaac acaaatatcg ttttccatac 2300

agcagtgcct atatagtgac tgattttaac tttcaatgtc catctttcaa 2350

aggaagtaac accaaggtac aatgttaaag gaatattcac tttacctagc 2400

agggaaaaat acacaaaaac tgcagatact tcatatagcc cattttaact 2450

tgtataaact gtgtgacttg tggcgtctta taaataatgc actgtaaaga 2500

ttactgaata gttgtgtcat gttaatgtgc ctaatttcat gtatcttgta 2550

atcatgattg agcctcagaa tcatttggag aaactatatt ttaaagaaca 2600

agacatactt caatgtatta tacagataaa gtattacatg tgtttgattt 2650

taaaagggcg gacattttat taaaatcaat attgtttttg ctttttctga 2700

ggagtctctt tcagtttcat tttttctcat cccatgactt ccctccgatg 2750

gt 2752

<210> 74
<211> 1001
<212> DNA
<213> Homo sapiens

<400> 74
agtatgaggc gcccaccatg gtgcgatgcg accgcgggct gcagatgctg 50

ctgaccacgg ccggagcctt cgccgccttc tcgctcatgg ccatcgccat 100

cggcaccgac tactggctgt actccagcgc gcacatctgc aacggcacca 150

acctgaccat ggacgacggg cccccgcccc gccgcgcccg cggcgacctc 200

acccactctg gtctgtggcg ggtgtgctgc atcgaaggga tctataaagg 250

gcactgcttc cggatcaatc acttcccaga ggacaatgac tacgaccacg 300

acagctcgga gtacctcctc cgcatcgtgc gagcctccag cgtcttcccc 350

atcctcagca ccatcctgct cctgctgggt ggcctgtgca tcggtgctgg 400

caggatctac agccgcaaga acaacatcgt cctcagtgcc ggcatcctct 450

tcgtggctgc aggcctcagt aacatcatcg gtatcatcgt ctacatttcc 500

agcaacacag gtgacccgag tgacaagcgg gacgaagaca aaaagaacca 550

ttacaactac ggctggtctt tttactttgg agctctgtct ttcattgtgg 600

ctgagaccgt gggcgtcctg gctgtaaaca tttacattga gaaaaataaa 650

gagttgaggt ttaagaccaa acgggaattc cttaaggcgt cttcctcttc 700

tccttatgcc aggatgccga gctacaggta ccggcgacgg cgctcgaggt 750

ccagctcaag gtccaccgag gcctcgccct ccagggacgt gtcgcccatg 800

ggcctgaaga tcacaggggc catccccatg ggggagctgt ccatgtacac 850

gctgtccagg gagcccctca aggtgaccac cgcagccagc tacagccccg 900

accaggaggc cagcttcctg caggtgcatg acttttttcca gcaggacctg 950

aaggaaggtt tccacgtcag catgctgaac cgacggacga cccctgtgtg 1000

a 1001

<210> 75
<211> 2805
<212> DNA
<213> Homo sapiens

<400> 75
cgggtctgat agtccctacc tgtcaggact ggtgttagga tgagataatg 50

tttgtgaact gtaaacatat ataaacgtgt gctactgtga gaactggaac 100

aaagaagaga gggagtgaga gaaatcaagg gagggctggg gctgggaaag 150

aacgaaaagg gagtcgcgta tagaggagag gcgacagtcg cgagccacac 200

tttgcaatga aactctttag actttctgcc gggagagcgg cccagacgcg 250

ccaggtctgt agcaggaggc cgcgagggcg ggtccccaga agcctacagg 300

tgagtatcgg ttctcccctt cccggctttc ggtccggagg aggcgggagc 350

agcttccctg ttctgatcct atcgcgggcg gcgcagggcc ggcttggcct 400

tccgtgggac ggggaggggg gcgggatgtg tcacccaaat accagtgggg 450

acggtcggtg gtggaaccag ccgggcaggt cgggtagagt ataagagccg 500

gagggagcgg ccggggcgca gacgcctgca gaccatccca gacgccggag 550

cccgagcccc gacgagtccc cgcgcctcat ccgcccgcgt ccggtccgcg 600

ttcctccgcc ccaccatggc tcggggcccc ggcctcgcgc cgccaccgct 650

gcggctgccg ctgctgctgc tggtgctggc ggcggtgacc ggccacacgg 700

ccgcgcagga caactgcacg tgtcccacca acaagatgac cgtgtgcagc 750

cccgacggcc ccggcggccg ctgccagtgc cgcgcgctgg gctcgggcat 800

ggcggtcgac tgctccacgc tgacctccaa gtgtctgctg ctcaaggcgc 850

gcatgagcgc ccccaagaac gcccgcacgc tggtgcggcc gagtgagcac 900

gcgctcgtgg acaacgatgg cctctacgac cccgactgcg accccgaggg 950

ccgcttcaag gcgcgccagt gcaaccagac gtcggtgtgc tggtgcgtga 1000

actcggtggg cgtgcgccgc acggacaagg gcgacctgag cctacgctgc 1050

gatgagctgg tgcgcaccca ccacatcctc attgacctgc gccaccgccc 1100

caccgccggc gccttcaacc actcagacct ggacgccgag ctgaggcggc 1150

tcttccgcga gcgctatcgg ctgcacccca agttcgtggc ggccgtgcac 1200

tacgagcagc ccaccatcca gatcgagctg cggcagaaca cgtctcagaa 1250

ggccgccggt gaagtggata tcggcgatgc cgcctactac ttcgagaggg 1300

acatcaaggg cgagtctcta ttccagggcc gcggcggcct ggacttgcgc 1350

gtgcgcggag aacccctgca ggtggagcgc acgctcatct attacctgga 1400

cgagattccc ccgaagttct ccatgaagcg cctcaccgcc ggcctcatcg 1450

ccgtcatcgt ggtggtcgtg gtggccctcg tcgccggcat ggccgtcctg 1500

gtgatcacca accggagaaa gtcggggaag tacaagaagg tggagatcaa 1550

ggaactgggg gagttgagaa aggaaccgag cttgtaggta cccggcgggg 1600

caggggatgg ggtggggtac cggatttcgg tatcgtccca gacccaagtg 1650

agtcacgctt cctgattcct cggcgcaaag gagacgttta tcctttcaaa 1700

ttcctgcctt ccccctccct tttgcgcaca caccaggttt aatagatcct 1750

ggcctcaggg tctccctttct ttctcacttc tgtcttgagg gaagcatttc 1800

taaaatgtat ccccttttcgg tccaacaaca ggaaacctga ctggggcagt 1850

gaaggaaggg atggcacagc gttatgtgta aaaaacaagt atctgtatga 1900

caacccggga tcgtttgcaa gtaactgaat ccattgcgac attgtgaagg 1950

cttaaatgag tttagatggg aaatagcgtt gttatcgcct tgggtttaaa 2000

ttatttgatg agttccactt gtatcatggc ctacccgagg agaagaggag 2050

tttgttaact gggcctatgt agtagcctca tttaccatcg tttgtattac 2100

tgaccacata tgcttgtcac tgggaaagaa gcctgtttca gctgcctgaa 2150

cgcagtttgg atgtctttga ggacagacat tgcccggaaa ctcagtctat 2200

ttattcttca gcttgccctt actgccactg atattggtaa tgttcttttt 2250

tgtaaaatgt ttgtacatat gttgtctttg ataatgttgc tgtaattttt 2300

taaaataaaa cacgaattta ataaaatatg ggaaaggcac aaaccagaag 2350

tcggcatttg tgaaaagtcc ctccagattt ctatcacttt ggtctctaat 2400

ttcccaagac ttgtattttt tttttatttc aaattataac actttttttt 2450

cccccagaag tgggtgtttc atgttgctac tctggtgtgt cccaagatat 2500

cctaactggc cagtgtaaat gctattcttt ctaaataaga ttatttggaa 2550

acttccttca aactgcagga gggcgagctc tgagggcacg agaagctaaa 2600

actagctgct tttgatgaaa aagagtgcca gtctttggtc atctctaaac 2650

aaggcttatc accaatggag acagaaaact ctagttcaag agctgtacct 2700

cctttgaatc ccagccctac tcgaaataag tggtactatt tccatttagc 2750

ctttgagcaa atcacttaac tcaaaggcgt tgtggctcta agattaaacg 2800

acttt 2805

<210> 76
<211> 2637
<212> DNA
<213> Homo sapiens

<400> 76
gcggccgcgg agcaagaagg gcgccgcgtc gtgcggcccg cgcagccccc 50

ggagccatgg gcaagtgcag cgggcgctgc acgctggtcg ccttctgctg 100

cctgcagctg gtggctgcgc tggagcggca gatctttgac ttcctgggct 150

accagtgggc tcccatccta gccaacttcc tgcacatcat ggcagtcatc 200

ctgggcatct ttggcaccgt gcagtaccgc tcccggtacc tcatcctgta 250

tgcagcctgg ctggtgctct gggttggctg gaatgcattt atcatctgct 300

tctacttgga ggttggacag ctgtcccagg accgggactt catcatgacc 350

ttcaacacat ccctgcaccg ctcctggtgg atggagaatg ggccaggctg 400

cctggtgaca cctgttctga actcccgcct ggctctggag gaccaccatg 450

tcatctctgt cactggctgc ctgcttgact acccctacat tgaagccctc 500

agcagcgccc tgcagatctt cctggcactg ttcggcttcg tgttcgcctg 550

ctacgtgagc aaagtgttcc tggaggagga ggacagcttt gacttcatcg 600

gcggctttga ctcctacgga taccaggcgc cccagaagac gtcgcattta 650

cagctgcagc ctctgtacac gtcggggtag cctctgcccc gcgcccaccc 700

cggcgcctcg ccctgggctg accgcagctg ccgcgagctc gggccaaggc 750

gcaggcgtgt ccccctggtg gcccgcgctc tcactgcagc ctgtgcccaa 800

ccccgcgtct gcatctggag atgcggactt ggacgtggac ttggacttgg 850

acttggattt gagcttggct cttcgcagcc cggacttcgg aggagtgggg 900

cggggcgggg gaggggcacc acgggttttt tgtttttttgt ttgtttgttt 950

ttaatctcag ccttggcgtg agctggggcc ttcctctctt ctccagcctc 1000

tccctttcac tcttcgccca gcatcctgcc cccctgtcca aaaacagcag 1050

gacatcaggc ccatcccatc ccaccacact cactcaccag ctctggggaa 1100

agctactgtg aactaggagc aggattcctg ggttctaatc gcaggtccat 1150

cactgactgt gacgtctagc aaagcccttg ccctctctga gcctcggttt 1200

ccgcacctca agtaattaat cccttagcaa atggactctt ttagacttct 1250

catttaactc aattccctga gctagactgg gattaaaatt ctcattttgc 1300

agtacattaa aactgatgcc cagagatgtg atttgcttga ggccacacag 1350

ctagattttt ggtggaagtg ggccttgaac acagtgtact ttctgcagtt 1400

tctgactgta aaacccagtg tctgctctct gagttccatt tccaagcccc 1450

cctccatctt ggacctatgt ggtctccacc atattcacac accaccacca 1500

ccacttgcca atgcctctct taaagcaata tacccattcg ttctcttatt 1550

gggaactgga tggatgaagc cccaaattca gccccaccca cagagaagcc 1600

ttcctacact cagcctctgt ccacccttgg caaatctttc aagctctctc 1650

ctccaggaaa gtggggcccc aactcagtca ctccaccccc ttccaggtcc 1700

ctgaggctgg ttctactgta tccccatcac ctccacaact ccactcaccc 1750

ctgacggctc catccacctc accagttgga aggcttgtgg tttcagagag 1800

gagcaatgct ggtcagcgct gcccagactc cagtgtttac agatcaccag 1850

catttacaac caatccaatg gccagaagcc tcctctaaca agcccagaag 1900

gagttctgaa ggggcagatg ggggtgtgag tagtcgggga gtcgggattg 1950

ccagcaccct cacccttcct tgggggcaag tagaggtgag aacactttcc 2000

ccacctccct ccacagacac tcctgaggac gctgcatccc acgcactgcc 2050

tggtgcgtcc atagagagag gatcaggtct cagcatttca tctgtgaaag 2100

aggcatggcc ctgggttaga aaggagggca ggagacatgg aggaactggg 2150

gggcacccag atggtgcaga tggtttgcac acctgagcct gtctgtggtg 2200

accattccgc tcctctccca ctaccctcca atctatcatt ccctactctc 2250

taaggccaaa atatcctgag caaggctggc aaccccaccc caccatccca 2300

aatgcaagca gccaggccca ggagttcctc tggcccccac aggcatggag 2350

ctcccagctg gtgggtacag cttgagaggg gggcagctcc ctcaggctaa 2400

gctactgccc ttcactgggc cagccctgcc tccagccctc acctctctca 2450

ccccaactct cccccaagcc cctttctact caacgggtgt agccactggt 2500

gctttgaagc cttttgtttt tataagatgg ttttgcaag gggaccaggt 2550

tctcttttca ctgggacctt gcaaggaggg gagtgctctc ctggtttctg 2600

tgcaggcggg ttgattaaag atggtgtttt cttctct 2637

<210> 77
<211> 1829
<212> DNA
<213> Homo sapiens

<400> 77
ggcggggaga ggagaggaga gaagagccgc ggggggccca gcccggagcc 50

aggatgcccg cgccgcgcgc ccgggagcag ccccgcgtgc ccggggagcg 100

ccagccgctg ctgcctcgcg gtgcgcgggg ccctcgacgg tggcggcggg 150

cggcgggcgc ggccgtgctg ctggtggaga tgctggagcg cgccgccttc 200

ttcggcgtca ccgccaacct cgtgctgtac ctcaacagca ccaacttcaa 250

ctggaccggc gagcaggcga cgcgcgccgc gctggtattc ctgggcgcct 300

cctacctgct ggcgcccgtg ggcggctggc tggccgacgt gtacctgggc 350

cgctaccgcg cggtcgcgct cagcctgctg ctctacctgg ccgcctcggg 400

cctgctgccc gccaccgcct tccccgacgg ccgcagctcc ttctgcggag 450

agatgcccgc gtcgccgctg ggacctgcct gcccctcggc cggctgcccg 500

cgctcctcgc ccagccccta ctgcgcgccc gtcctctacg cgggcctgct 550

gctactcggc ctggccgcca gctccgtccg gagcaacctc acctccttcg 600

gtgccgacca ggtgatggat ctcggccgcg acgccacccg ccgcttcttc 650

aactggtttt actggagcat caacctgggt gctgtgctgt cgctgctggt 700

ggtggcgttt attcagcaga acatcagctt cctgctgggc tacagcatcc 750

ctgtgggctg tgtgggcctg gcatttttca tcttcctctt tgccaccccc 800

gtcttcatca ccaagccccc gatgggcagc caagtgtcct ctatgcttaa 850

gctcgctctc caaaactgct gcccccagct gtggcaacga cactcggcca 900

gagaccgtca atgtgcccgc gtgctggccg acgagaggtc tccccagcca 950

ggggcttccc cgcaagagga catcgccaac ttccaggtgc tggtgaagat 1000

cttgcccgtc atggtgaccc tggtgcccta ctggatggtc tacttccaga 1050

tgcagtccac ctatgtcctg cagggtcttc acctccacat cccaaacatt 1100

ttcccagcca acccggccaa catctctgtg gccctgagag cccagggcag 1150

cagctacacg atcccggaag cctggctcct cctggccaat gttgtggtgg 1200

tgctgattct ggtccctctg aaggaccgct tgatcgaccc tttactgctg 1250

cggtgcaagc tgcttccctc tgctctgcag aagatggcgc tggggatgtt 1300

ctttggtttt acctccgtca ttgtggcagg agtcctggag atggagcgct 1350

tacactacat ccaccacaac gagaccgtgt cccagcagat tggggaggtc 1400

ctgtacaacg cggcaccact gtccatctgg tggcagatcc ctcagtacct 1450

gctcattggg atcagtgaga tctttgccag catcccaggc ctggagtttg 1500

cctactcaga ggccccgcgc tccatgcagg gcgccatcat gggcatcttc 1550

ttctgcctgt cgggggtggg ctcactgttg ggctccagcc tagtggcact 1600

gctgtccttg cccgggggct ggctgcactg ccccaaggac tttgggaaca 1650

tcaacaattg ccggatggac ctctacttct tcctgctggc tggcattcag 1700

gccgtcacgg ctctcctatt tgtctggatc gctggacgct atgagagggc 1750

gtcccagggc ccagcctccc acagccgttt cagcagggac aggggctgaa 1800

caggccctat tccagccccc ttgcttcac 1829

<210> 78
<211> 5746
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 5621, 5664
<223> unknown base

<400> 78
ggccgcgggc tctcgcgggg cggcgacgcc gcggggagga tgctgcttgc 50

cgcgcccgcg tcctcaccgt cctcccgggc cgcctgctgg ggctttgttg 100

tggcccggac gccgcgggcc acccccctgaa gtcgcctgcc gccgccgccg 150

ccgcacctag cggacgggcg ggcgggcgcg cgtgtgccca ggagtgcgcg 200

cctgtcgcgg tggtgggtgc aggactggac ccacgggccc attgtgcgcc 250

cgcccgcggc agccaggacc atgtgggtga acccggagga ggtgttgctg 300

gccaacgcgc tgtggatcac cgagagggcc aacccatact tcatcctgca 350

gcggaggaag ggccacgccg gcgatggagg cggcggcggc ggactggcgg 400

gcctgctggt gggtaccctt gatgttgtgt tggactccag cgcccgggtc 450

gctccttacc gaatcttgta ccagactcca gactccctgg tctactggac 500

catcgcctgt ggtggttcca ggaaagaaat cactgaacac tgggaatggc 550

ttgagcaaaa tctcttgcag acactctcca tctttgaaaa tgagaatgat 600

atcaccacat ttgtgagagg aaaaatacag ggcatcattg cagaatacaa 650

caaaatcaat gatgtaaagg aagatgatga cacggagaag tttaaagaag 700

ccattgtgaa atttcatagg ctgtttggga tgccagagga agagaaactc 750

gtcaactatt actcttgcag ctattggaag gggaaggtcc cccgtcaggg 800

ttggatgtac ctcagcatta accacctttg cttttattct tttcttatgg 850

gaagggaagc gaaactggtc atccggtggg tagacatcac tcagcttgag 900

aagaatgcca ccctgcttct gcctgatgtg atcaaagtga gcacacggtc 950

cagtgagcat ttcttctctg tattcctcaa catcaacgag accttcaagt 1000

taatggagca gcttgccaac atagccatga ggcaactctt agacaatgag 1050

ggatttgaac aagatcgatc cctgcccaaa ctcaaaagga aatctcctaa 1100

aaaagtgtct gctctaaaac gtgatcttga tgccagggca aagagtgaga 1150

gataccgtgc acttttccgg ctgcccaaag atgaaaaatt agatggccac 1200

acagactgca ctctctggac tccatttaac aaaatgcaca ttttggggca 1250

gatgtttgtg tccacaaatt acatctgttt taccagcaag gaggagaact 1300

tatgtagcct cattatcccg ctccgtgagg tgacaattgt ggaaaaggca 1350

gacagctcca gtgtgctccc cagtccctta tccatcagca cccgaaacag 1400

gatgaccttc ctatttgcca acttgaaaga tagagacttt ctagtgcaga 1450

ggatctcaga tttcctgcaa cagactactt ccaaaatata ttctgacaag 1500

gagtttgcag gaagttacaa cagttcagat gatgaggtgt actctcgacc 1550

cagcagcctc gtctcctcca gcccccagag aagcacgagc tctgatgctg 1600

atggagagcg ccagtttaac ctaaatggca acagcgtccc cacagccaca 1650

cagaccctga tgaccatgta tcggcggcgg tctcccgagg agttcaaccc 1700

gaaattggcc aaagagtttc tgaaagagca agcctggaag attcactttg 1750

ctgagtatgg gcaagggatc tgcatgtacc gcacagagaa aacgcgggag 1800

ctggtgttga agggcatccc ggagagcatg cgtggggagc tctggctgct 1850

gctgtcaggt gccatcaatg agaaggccac acatcctggg tactatgaag 1900

acctagtgga gaagtccatg gggaagtata atctcgccac ggaggagatt 1950

gagagggatt tacaccgctc ccttccagaa cacccagctt ttcagaatga 2000

aatgggcatt gctgcactaa ggagagtctt aacagcttat gcttttcgaa 2050

atcccaacat agggtattgc caggccatga atattgtcac ttcagtgctg 2100

ctgctttatg ccaaagagga ggaagctttc tggctgcttg tggctttgtg 2150

tgagcgcatg ctcccagatt actacaacac cagagttgtg ggtgcactgg 2200

tggaccaagg tgtctttgag gagctagcac gagactacgt cccacagctg 2250

tacgactgca tgcaagacct gggcgtgatt tccaccatct ccctgtcttg 2300

gttcctcaca ctatttctca gtgtgatgcc ttttgagagt gcagttgtgg 2350

ttgttgactg tttcttctat gaaggaatta aagtgatatt ccagttggcc 2400

ctagctgtgc tggatgcaaa tgtggacaaa ctgttgaact gcaaggatga 2450

tggggaggcc atgaccgttt tgggaaggta tttagacagt gtgaccaata 2500

aagacagcac actgcctccc attcctcacc tccactcctt gctcagcgat 2550

gatgtggaac cttaccctga ggtagacatc tttagactca tcagaacttc 2600

ctacgagaaa ttcggaacta tccgggcaga tttgattgaa cagatgagat 2650

tcaaacagag actgaaagtg atccagacgc tggaggatac tacgaaacgc 2700

aacgtggtac gaaccattgt gacagaaact tcctttacca ttgatgagct 2750

ggaagaactt tatgctcttt tcaaggcaga acatctcacc agctgctact 2800

ggggcgggag cagcaacgcg ctggaccggc atgaccccag cctgccctac 2850

ctggaacagt atcgcattga cttcgagcag ttcaagggaa tgtttgctct 2900

tctctttcct tgggcatgtg gaactcactc tgacgttctg gcctcccgct 2950

tgttccagtt attagatgaa aatggagact ctttgattaa cttccgggag 3000

tttgtctctg ggctaagtgc tgcatgccat ggggacctca cagagaagct 3050

caaactcctg tacaaaatgc acgtcttgcc tgagccatcc tctgatcaag 3100

atgaaccaga ttctgctttt gaagcaactc agtacttctt tgaagatatt 3150

accccagaat gtacacatgt tgttggattg gatagcagaa gcaaacaggg 3200

tgcagatgat ggctttgtta cggtgagcct aaagccagac aaagggaaga 3250

gagcaaattc ccaagaaaat cgtaattatt tgagactgtg gactccagaa 3300

aataaatcta agtcaaagaa tgcaaaggat ttacccaaat taaatcaggg 3350

gcagttcatt gaactgtgta agacaatgta taacatgttc agcgaagacc 3400

ccaatgagca ggagctgtac catgccacgg cagcagtgac cagcctcctg 3450

ctggagattg gggaggtcgg caagttgttc gtggcccagc ctgcaaagga 3500

gggcgggagc ggaggcagtg ggccgtcctg ccaccagggc atcccaggcg 3550

tgctcttccc caagaaaggg ccaggccagc cttacgtggt ggagtctgtt 3600

gagcccctgc cggccagcct ggcccccgac agcgaggaac actcccttgg 3650

aggacaaatg gaggacatca agctggagga ctcctcgccc cgggacaacg 3700

gggcctgctc ctccatgctg atctctgacg acgacaccaa ggacgacagc 3750

tccatgtcct catactcggt gctgagtgcc ggctcccacg aggaggacaa 3800

gctgcactgc gaggacatcg gagaggacac ggtcctggtg cggagcggcc 3850

agggcacggc ggcactgccc cggagcacca gcctggaccg ggactgggcc 3900

atcaccttcg agcagttcct ggcctccctc ttaactgagc ctgccctggt 3950

caagtacttt gacaagcccg tgtgcatgat ggccaggatt accagtgcaa 4000

aaaacatccg gatgatgggc aagcccctca cctcggccag tgactatgaa 4050

atctcggcca tgtccggctg acacgggcgc cttcccgggg gagtgggagg 4100

agagggaggg gagggatttt ttatgttctt ctgtgttgag ttttttcttt 4150

ctttctttta aattaaatat ttattagtac ctggcttgaa gcctagtgtt 4200

ttcataatgt aattcaatga aaactgttgg agaaatattt aaacacctca 4250

atgtaggtac attacactct tgttgcgggg aggggattta ccagaataca 4300

gtttatttcg tgaattctaa aaaacaaaaa gatgaatctg tcagtgatat 4350

gtgtgtatta taacttatta atcttgctgt tgagctgtat acatggttta 4400

aaaaatagta ctgtttaatg ctaagtaagg cagcagtcat ttgtgtattc 4450

aggcttttta aataaaatta gagctgtaag gaaaatgaaa agccacaaat 4500

gcaagactgt tcttaaatgg aaggcatagt cagcgagggt aaatcctata 4550

ccactttagg aagtattaaa aatattttta agatttgaaa tatatttcat 4600

agaagtcctc tattcaaaat catattccac agatgttccc cttcaaaggg 4650

aaaacatttg gggttctaaa cagttatgaa agtaagtgat ttttacatga 4700

ttccagaata acacttgtat tgaccaattt agacagatac cagaccaatt 4750

ttgcatttaa gaaattgttc tgattattta cgtcaactca ttagaattca 4800

gtgaaaagta acagtctttt gtcacagaga atctgaaagt agcagcaaag 4850

acagagggct catgacaggt ttttgctttt gctttgcttt tgtttttgaa 4900

agagtaaaag tactgatgct tctgatactg gatgtttagc ttcttactgc 4950

aaaaacataa gtaaaacagt caactttacc atttccgtat tctccataga 5000

ttgaagaaat ttataccaca tatcgcatat gaccatcttt ccatcaaatc 5050

aatgtagaga taatgtaaac tgaaaaaaaa tctgcaagat aatgtaactg 5100

aatgttttaa aaacagaact tgtcacttta tataaaagaa tagtatgctc 5150

tatttcctga atggatgtgg aaatgaaagc tagcgcacct gcactttgaa 5200

ttcttgcttc ttttttatta ctgttatgat tttgcttttt acagatgttg 5250

gacgattttt tcttctgatt gttgaattca taatcatggt ctcatttcct 5300

ttgcttcttt ggaatatttc tttcaacaca ttcctttatt ttattataca 5350

ttgtgtcctt tttttagcta ttgctgctgt tgtttttttat tctatttaca 5400

ggatgatttt taaactgtca aatgaagtag tgttaacctc aaataggcta 5450

aatgtgaaca aataaaatac agcaaatact cagatacagc tttttatctt 5500

tgtgcttgag ttcctgccta aggaataaca ttattctttt gacaactttt 5550

gcaggggaaa ttatatcagg caaccatttt gattaagtaa ataaatttta 5600

taggcaaaca tatagagaga natacaattt gtagtatatc aatgactata 5650

tttaaaataa ggantataat tgttatcagt tatctaactt aaaatgctta 5700

tccataatga tcagtgatat tcagcttttt aaaatatgct tgttgg 5746

<210> 79
<211> 183
<212> PRT
<213> Homo sapiens

<400> 79
Met Arg Ala Ala Tyr Leu Phe Leu Leu Phe Leu Pro Ala Gly Leu
1 5 10 15

Leu Ala Gln Gly Gln Tyr Asp Leu Asp Pro Leu Pro Pro Phe Pro
20 25 30

Asp His Val Gln Tyr Thr His Tyr Ser Asp Gln Ile Asp Asn Pro
35 40 45

Asp Tyr Tyr Asp Tyr Gln Glu Val Thr Pro Arg Pro Ser Glu Glu
50 55 60

Gln Phe Gln Phe Gln Ser Gln Gln Gln Val Gln Gln Glu Val Ile
65 70 75

Pro Ala Pro Thr Pro Glu Pro Gly Asn Ala Glu Leu Glu Pro Thr
80 85 90

Glu Pro Gly Pro Leu Asp Cys Arg Glu Glu Gln Tyr Pro Cys Thr
95 100 105

Arg Leu Tyr Ser Ile His Arg Pro Cys Lys Gln Cys Leu Asn Glu
110 115 120

Val Cys Phe Tyr Ser Leu Arg Arg Val Tyr Val Ile Asn Lys Glu
125 130 135

Ile Cys Val Arg Thr Val Cys Ala His Glu Glu Leu Leu Arg Ala
140 145 150

Asp Leu Cys Arg Asp Lys Phe Ser Lys Cys Gly Val Met Ala Ser
155 160 165

Ser Gly Leu Cys Gln Ser Val Ala Ala Ser Cys Ala Arg Ser Cys
170 175 180

Gly Ser Cys

<210> 80
<211> 909
<212> PRT
<213> Homo sapiens

<400> 80

```
Met Ala Ala Arg Pro Leu Pro Val Ser Pro Ala Arg Ala Leu Leu
  1               5                  10                  15

Leu Ala Leu Ala Gly Ala Leu Leu Ala Pro Cys Glu Ala Arg Gly
                20                  25                  30

Val Ser Leu Trp Asn Glu Gly Arg Ala Asp Glu Val Val Ser Ala
                35                  40                  45

Ser Val Arg Ser Gly Asp Leu Trp Ile Pro Val Lys Ser Phe Asp
                50                  55                  60

Ser Lys Asn His Pro Glu Val Leu Asn Ile Arg Leu Gln Arg Glu
                65                  70                  75

Ser Lys Glu Leu Ile Ile Asn Leu Glu Arg Asn Glu Gly Leu Ile
                80                  85                  90

Ala Ser Ser Phe Thr Glu Thr His Tyr Leu Gln Asp Gly Thr Asp
                95                 100                 105

Val Ser Leu Ala Arg Asn Tyr Thr Val Ile Leu Gly His Cys Tyr
               110                 115                 120

Tyr His Gly His Val Arg Gly Tyr Ser Asp Ser Ala Val Ser Leu
               125                 130                 135

Ser Thr Cys Ser Gly Leu Arg Gly Leu Ile Val Phe Glu Asn Glu
               140                 145                 150

Ser Tyr Val Leu Glu Pro Met Lys Ser Ala Thr Asn Arg Tyr Lys
               155                 160                 165

Leu Phe Pro Ala Lys Lys Leu Lys Ser Val Arg Gly Ser Cys Gly
               170                 175                 180

Ser His His Asn Thr Pro Asn Leu Ala Ala Lys Asn Val Phe Pro
               185                 190                 195

Pro Pro Ser Gln Thr Trp Ala Arg Arg His Lys Arg Glu Thr Leu
               200                 205                 210

Lys Ala Thr Lys Tyr Val Glu Leu Val Ile Val Ala Asp Asn Arg
               215                 220                 225

Glu Phe Gln Arg Gln Gly Lys Asp Leu Glu Lys Val Lys Gln Arg
               230                 235                 240

Leu Ile Glu Ile Ala Asn His Val Asp Lys Phe Tyr Arg Pro Leu
               245                 250                 255
```

```
Asn Ile Arg Ile Val Leu Val Gly Val Glu Val Trp Asn Asp Met
                260                 265                 270

Asp Lys Cys Ser Val Ser Gln Asp Pro Phe Thr Ser Leu His Glu
                275                 280                 285

Phe Leu Asp Trp Arg Lys Met Lys Leu Leu Pro Arg Lys Ser His
                290                 295                 300

Asp Asn Ala Gln Leu Val Ser Gly Val Tyr Phe Gln Gly Thr Thr
                305                 310                 315

Ile Gly Met Ala Pro Ile Met Ser Met Cys Thr Ala Asp Gln Ser
                320                 325                 330

Gly Gly Ile Val Met Asp His Ser Asp Asn Pro Leu Gly Ala Ala
                335                 340                 345

Val Thr Leu Ala His Glu Leu Gly His Asn Phe Gly Met Asn His
                350                 355                 360

Asp Thr Leu Asp Arg Gly Cys Ser Cys Gln Met Ala Val Glu Lys
                365                 370                 375

Gly Gly Cys Ile Met Asn Ala Ser Thr Gly Tyr Pro Phe Pro Met
                380                 385                 390

Val Phe Ser Ser Cys Ser Arg Lys Asp Leu Glu Thr Ser Leu Glu
                395                 400                 405

Lys Gly Met Gly Val Cys Leu Phe Asn Leu Pro Glu Val Arg Glu
                410                 415                 420

Ser Phe Gly Gly Gln Lys Cys Gly Asn Arg Phe Val Glu Glu Gly
                425                 430                 435

Glu Glu Cys Asp Cys Gly Glu Pro Glu Glu Cys Met Asn Arg Cys
                440                 445                 450

Cys Asn Ala Thr Thr Cys Thr Leu Lys Pro Asp Ala Val Cys Ala
                455                 460                 465

His Gly Leu Cys Cys Glu Asp Cys Gln Leu Lys Pro Ala Gly Thr
                470                 475                 480

Ala Cys Arg Asp Ser Ser Asn Ser Cys Asp Leu Pro Glu Phe Cys
                485                 490                 495

Thr Gly Ala Ser Pro His Cys Pro Ala Asn Val Tyr Leu His Asp
                500                 505                 510

Gly His Ser Cys Gln Asp Val Asp Gly Tyr Cys Tyr Asn Gly Ile
                515                 520                 525

Cys Gln Thr His Glu Gln Gln Cys Val Thr Leu Trp Gly Pro Gly
                530                 535                 540

Ala Lys Pro Ala Pro Gly Ile Cys Phe Glu Arg Val Asn Ser Ala
                545                 550                 555

Gly Asp Pro Tyr Gly Asn Cys Gly Lys Val Ser Lys Ser Ser Phe
```

```
                560                 565                 570

Ala Lys Cys Glu Met Arg Asp Ala Lys Cys Gly Lys Ile Gln Cys
                575                 580                 585

Gln Gly Gly Ala Ser Arg Pro Val Ile Gly Thr Asn Ala Val Ser
                590                 595                 600

Ile Glu Thr Asn Ile Pro Leu Gln Gln Gly Gly Arg Ile Leu Cys
                605                 610                 615

Arg Gly Thr His Val Tyr Leu Gly Asp Asp Met Pro Asp Pro Gly
                620                 625                 630

Leu Val Leu Ala Gly Thr Lys Cys Ala Asp Gly Lys Ile Cys Leu
                635                 640                 645

Asn Arg Gln Cys Gln Asn Ile Ser Val Phe Gly Val His Glu Cys
                650                 655                 660

Ala Met Gln Cys His Gly Arg Gly Val Cys Asn Asn Arg Lys Asn
                665                 670                 675

Cys His Cys Glu Ala His Trp Ala Pro Pro Phe Cys Asp Lys Phe
                680                 685                 690

Gly Phe Gly Gly Ser Thr Asp Ser Gly Pro Ile Arg Gln Ala Asp
                695                 700                 705

Asn Gln Gly Leu Thr Ile Gly Ile Leu Val Thr Ile Leu Cys Leu
                710                 715                 720

Leu Ala Ala Gly Phe Val Val Tyr Leu Lys Arg Lys Thr Leu Ile
                725                 730                 735

Arg Leu Leu Phe Thr Asn Lys Lys Thr Thr Ile Glu Lys Leu Arg
                740                 745                 750

Cys Val Arg Pro Ser Arg Pro Pro Arg Gly Phe Gln Pro Cys Gln
                755                 760                 765

Ala His Leu Gly His Leu Gly Lys Gly Leu Met Arg Lys Pro Pro
                770                 775                 780

Asp Ser Tyr Pro Pro Lys Asp Asn Pro Arg Arg Leu Leu Gln Cys
                785                 790                 795

Gln Asn Val Asp Ile Ser Arg Pro Leu Asn Gly Leu Asn Val Pro
                800                 805                 810

Gln Pro Gln Ser Thr Gln Arg Val Leu Pro Pro Leu His Arg Ala
                815                 820                 825

Pro Arg Ala Pro Ser Val Pro Ala Arg Pro Leu Pro Ala Lys Pro
                830                 835                 840

Ala Leu Arg Gln Ala Gln Gly Thr Cys Lys Pro Asn Pro Pro Gln
                845                 850                 855

Lys Pro Leu Pro Ala Asp Pro Leu Ala Arg Thr Thr Arg Leu Thr
                860                 865                 870
```

```
 His Ala Leu Ala Arg Thr Pro Gly Gln Trp Glu Thr Gly Leu Arg
                 875                 880                 885

 Leu Ala Pro Leu Arg Pro Ala Pro Gln Tyr Pro His Gln Val Pro
                 890                 895                 900

 Arg Ser Thr His Thr Ala Tyr Ile Lys
                 905

<210> 81
<211> 158
<212> PRT
<213> Homo sapiens

<400> 81
 Met Ala Ser Arg Ser Met Arg Leu Leu Leu Leu Leu Ser Cys Leu
   1               5                  10                  15

 Ala Lys Thr Gly Val Leu Gly Asp Ile Ile Met Arg Pro Ser Cys
                 20                  25                  30

 Ala Pro Gly Trp Phe Tyr His Lys Ser Asn Cys Tyr Gly Tyr Phe
                 35                  40                  45

 Arg Lys Leu Arg Asn Trp Ser Asp Ala Glu Leu Glu Cys Gln Ser
                 50                  55                  60

 Tyr Gly Asn Gly Ala His Leu Ala Ser Ile Leu Ser Leu Lys Glu
                 65                  70                  75

 Ala Ser Thr Ile Ala Glu Tyr Ile Ser Gly Tyr Gln Arg Ser Gln
                 80                  85                  90

 Pro Ile Trp Ile Gly Leu His Asp Pro Gln Lys Arg Gln Gln Trp
                 95                 100                 105

 Gln Trp Ile Asp Gly Ala Met Tyr Leu Tyr Arg Ser Trp Ser Gly
                110                 115                 120

 Lys Ser Met Gly Gly Asn Lys His Cys Ala Glu Ile Ser Ser Asn
                125                 130                 135

 Asn Asn Phe Leu Thr Trp Ser Ser Asn Glu Cys Asn Lys Arg Gln
                140                 145                 150

 His Phe Leu Cys Lys Tyr Arg Pro
                155

<210> 82
<211> 84
<212> PRT
<213> Homo sapiens

<400> 82
 Met Ala Thr Met Glu Asn Lys Val Ile Cys Ala Leu Val Leu Val
   1               5                  10                  15

 Ser Met Leu Ala Leu Gly Thr Leu Ala Glu Ala Gln Thr Glu Thr
                 20                  25                  30
```

```
 Cys Thr Val Ala Pro Arg Glu Arg Gln Asn Cys Gly Phe Pro Gly
                  35                  40                  45

 Val Thr Pro Ser Gln Cys Ala Asn Lys Gly Cys Cys Phe Asp Asp
                  50                  55                  60

 Thr Val Arg Gly Val Pro Trp Cys Phe Tyr Pro Asn Thr Ile Asp
                  65                  70                  75

 Val Pro Pro Glu Glu Glu Cys Glu Phe
                  80

<210> 83
<211> 680
<212> PRT
<213> Homo sapiens

<400> 83
 Met Leu Pro Gln Ile Pro Phe Leu Leu Leu Val Ser Leu Asn Leu
   1               5                  10                  15

 Val His Gly Val Phe Tyr Ala Glu Arg Tyr Gln Thr Pro Thr Gly
                  20                  25                  30

 Ile Lys Gly Pro Leu Pro Asn Thr Lys Thr Gln Phe Phe Ile Pro
                  35                  40                  45

 Tyr Thr Ile Lys Ser Lys Gly Ile Ala Val Arg Gly Glu Gln Gly
                  50                  55                  60

 Thr Pro Gly Pro Pro Gly Pro Ala Gly Pro Arg Gly His Pro Gly
                  65                  70                  75

 Pro Ser Gly Pro Pro Gly Lys Pro Gly Tyr Gly Ser Pro Gly Leu
                  80                  85                  90

 Gln Gly Glu Pro Gly Leu Pro Gly Pro Pro Gly Pro Ser Ala Val
                  95                 100                 105

 Gly Lys Pro Gly Val Pro Gly Leu Pro Gly Lys Pro Gly Glu Arg
                 110                 115                 120

 Gly Pro Tyr Gly Pro Lys Gly Asp Val Gly Pro Ala Gly Leu Pro
                 125                 130                 135

 Gly Pro Arg Gly Pro Pro Gly Pro Pro Gly Ile Pro Gly Pro Ala
                 140                 145                 150

 Gly Ile Ser Val Pro Gly Lys Pro Gly Gln Gln Gly Pro Thr Gly
                 155                 160                 165

 Ala Pro Gly Pro Arg Gly Phe Pro Gly Glu Lys Gly Ala Pro Gly
                 170                 175                 180

 Val Pro Gly Met Asn Gly Gln Lys Gly Glu Met Gly Tyr Gly Ala
                 185                 190                 195

 Pro Gly Arg Pro Gly Glu Arg Gly Leu Pro Gly Pro Gln Gly Pro
                 200                 205                 210

 Thr Gly Pro Ser Gly Pro Pro Gly Val Gly Lys Arg Gly Glu Asn
```

```
                215                 220                 225

Gly Val Pro Gly Gln Pro Gly Ile Lys Gly Asp Arg Gly Phe Pro
                230                 235                 240

Gly Glu Met Gly Pro Ile Gly Pro Pro Gly Pro Gln Gly Pro Pro
                245                 250                 255

Gly Glu Arg Gly Pro Glu Gly Ile Gly Lys Pro Gly Ala Ala Gly
                260                 265                 270

Ala Pro Gly Gln Pro Gly Ile Pro Gly Thr Lys Gly Leu Pro Gly
                275                 280                 285

Ala Pro Gly Ile Ala Gly Pro Pro Gly Pro Pro Gly Phe Gly Lys
                290                 295                 300

Pro Gly Leu Pro Gly Leu Lys Gly Glu Arg Gly Pro Ala Gly Leu
                305                 310                 315

Pro Gly Gly Pro Gly Ala Lys Gly Glu Gln Gly Pro Ala Gly Leu
                320                 325                 330

Pro Gly Lys Pro Gly Leu Thr Gly Pro Pro Gly Asn Met Gly Pro
                335                 340                 345

Gln Gly Pro Lys Gly Ile Pro Gly Ser His Gly Leu Pro Gly Pro
                350                 355                 360

Lys Gly Glu Thr Gly Pro Ala Gly Pro Ala Gly Tyr Pro Gly Ala
                365                 370                 375

Lys Gly Glu Arg Gly Ser Pro Gly Ser Asp Gly Lys Pro Gly Tyr
                380                 385                 390

Pro Gly Lys Pro Gly Leu Asp Gly Pro Lys Gly Asn Pro Gly Leu
                395                 400                 405

Pro Gly Pro Lys Gly Asp Pro Gly Val Gly Gly Pro Pro Gly Leu
                410                 415                 420

Pro Gly Pro Val Gly Pro Ala Gly Ala Lys Gly Met Pro Gly His
                425                 430                 435

Asn Gly Glu Ala Gly Pro Arg Gly Ala Pro Gly Ile Pro Gly Thr
                440                 445                 450

Arg Gly Pro Ile Gly Pro Pro Gly Ile Pro Gly Phe Pro Gly Ser
                455                 460                 465

Lys Gly Asp Pro Gly Ser Pro Gly Pro Pro Gly Pro Ala Gly Ile
                470                 475                 480

Ala Thr Lys Gly Leu Asn Gly Pro Thr Gly Pro Pro Gly Pro Pro
                485                 490                 495

Gly Pro Arg Gly Pro Ser Gly Glu Pro Gly Leu Pro Gly Pro Pro
                500                 505                 510

Gly Pro Pro Gly Pro Pro Gly Gln Ala Val Met Pro Glu Gly Phe
                515                 520                 525
```

```
 Ile Lys Ala Gly Gln Arg Pro Ser Leu Ser Gly Thr Pro Leu Val
                 530                 535                 540

 Ser Ala Asn Gln Gly Val Thr Gly Met Pro Val Ser Ala Phe Thr
                 545                 550                 555

 Val Ile Leu Ser Lys Ala Tyr Pro Ala Ile Gly Thr Pro Ile Pro
                 560                 565                 570

 Phe Asp Lys Ile Leu Tyr Asn Arg Gln Gln His Tyr Asp Pro Arg
                 575                 580                 585

 Thr Gly Ile Phe Thr Cys Gln Ile Pro Gly Ile Tyr Tyr Phe Ser
                 590                 595                 600

 Tyr His Val His Val Lys Gly Thr His Val Trp Val Gly Leu Tyr
                 605                 610                 615

 Lys Asn Gly Thr Pro Val Met Tyr Thr Tyr Asp Glu Tyr Thr Lys
                 620                 625                 630

 Gly Tyr Leu Asp Gln Ala Ser Gly Ser Ala Ile Ile Asp Leu Thr
                 635                 640                 645

 Glu Asn Asp Gln Val Trp Leu Gln Leu Pro Asn Ala Glu Ser Asn
                 650                 655                 660

 Gly Leu Tyr Ser Ser Glu Tyr Val His Ser Ser Phe Ser Gly Phe
                 665                 670                 675

 Leu Val Ala Pro Met
                 680

<210> 84
<211> 542
<212> PRT
<213> Homo sapiens

<400> 84
 Met Pro Leu Pro Pro Leu Leu Leu Leu Leu Leu Ala Ala Pro Trp
   1               5                  10                  15

 Gly Arg Ala Val Pro Cys Val Ser Gly Gly Leu Pro Lys Pro Ala
                  20                  25                  30

 Asn Ile Thr Phe Leu Ser Ile Asn Met Lys Asn Val Leu Gln Trp
                  35                  40                  45

 Thr Pro Pro Glu Gly Leu Gln Gly Val Lys Val Thr Tyr Thr Val
                  50                  55                  60

 Gln Tyr Phe Ile Tyr Gly Gln Lys Lys Trp Leu Asn Lys Ser Glu
                  65                  70                  75

 Cys Arg Asn Ile Asn Arg Thr Tyr Cys Asp Leu Ser Ala Glu Thr
                  80                  85                  90

 Ser Asp Tyr Glu His Gln Tyr Tyr Ala Lys Val Lys Ala Ile Trp
                  95                 100                 105
```

```
Gly Thr Lys Cys Ser Lys Trp Ala Glu Ser Gly Arg Phe Tyr Pro
                110                 115                 120

Phe Leu Glu Thr Gln Ile Gly Pro Pro Glu Val Ala Leu Thr Thr
                125                 130                 135

Asp Glu Lys Ser Ile Ser Val Val Leu Thr Ala Pro Glu Lys Trp
                140                 145                 150

Lys Arg Asn Pro Glu Asp Leu Pro Val Ser Met Gln Gln Ile Tyr
                155                 160                 165

Ser Asn Leu Lys Tyr Asn Val Ser Val Leu Asn Thr Lys Ser Asn
                170                 175                 180

Arg Thr Trp Ser Gln Cys Val Thr Asn His Thr Leu Val Leu Thr
                185                 190                 195

Trp Leu Glu Pro Asn Thr Leu Tyr Cys Val His Val Glu Ser Phe
                200                 205                 210

Val Pro Gly Pro Pro Arg Arg Ala Gln Pro Ser Glu Lys Gln Cys
                215                 220                 225

Ala Arg Thr Leu Lys Asp Gln Ser Ser Glu Phe Lys Ala Lys Ile
                230                 235                 240

Ile Phe Trp Tyr Val Leu Pro Ile Ser Ile Thr Val Phe Leu Phe
                245                 250                 255

Ser Val Met Gly Tyr Ser Ile Tyr Arg Tyr Ile His Val Gly Lys
                260                 265                 270

Glu Lys His Pro Ala Asn Leu Ile Leu Ile Tyr Gly Asn Glu Phe
                275                 280                 285

Asp Lys Arg Phe Phe Val Pro Ala Glu Lys Ile Val Ile Asn Phe
                290                 295                 300

Ile Thr Leu Asn Ile Ser Asp Asp Ser Lys Ile Ser His Gln Asp
                305                 310                 315

Met Ser Leu Leu Gly Lys Ser Ser Asp Val Ser Ser Leu Asn Asp
                320                 325                 330

Pro Gln Pro Ser Gly Asn Leu Arg Pro Pro Gln Glu Glu Glu Glu
                335                 340                 345

Val Lys His Leu Gly Tyr Ala Ser His Leu Met Glu Ile Phe Cys
                350                 355                 360

Asp Ser Glu Glu Asn Thr Glu Gly Thr Ser Leu Thr Gln Gln Glu
                365                 370                 375

Ser Leu Ser Arg Thr Ile Pro Pro Asp Lys Thr Val Ile Glu Tyr
                380                 385                 390

Glu Tyr Asp Val Arg Thr Thr Asp Ile Cys Ala Gly Pro Glu Glu
                395                 400                 405

Gln Glu Leu Ser Leu Gln Glu Glu Val Ser Thr Gln Gly Thr Leu
```

```
                410                 415                 420

 Leu Glu Ser Gln Ala Ala Leu Ala Val Leu Gly Pro Gln Thr Leu
                425                 430                 435

 Gln Tyr Ser Tyr Thr Pro Gln Leu Gln Asp Leu Asp Pro Leu Ala
                440                 445                 450

 Gln Glu His Thr Asp Ser Glu Glu Gly Pro Glu Glu Glu Pro Ser
                455                 460                 465

 Thr Thr Leu Val Asp Trp Asp Pro Gln Thr Gly Arg Leu Cys Ile
                470                 475                 480

 Pro Ser Leu Ser Ser Phe Asp Gln Asp Ser Glu Gly Cys Glu Pro
                485                 490                 495

 Ser Glu Gly Asp Gly Leu Gly Glu Glu Gly Leu Leu Ser Arg Leu
                500                 505                 510

 Tyr Glu Glu Pro Ala Pro Asp Arg Pro Pro Gly Glu Asn Glu Thr
                515                 520                 525

 Tyr Leu Met Gln Phe Met Glu Glu Trp Gly Leu Tyr Val Gln Met
                530                 535                 540

 Glu Asn


<210> 85
<211> 553
<212> PRT
<213> Homo sapiens

<400> 85
 Met Arg Ala Pro Gly Arg Pro Ala Leu Arg Pro Leu Pro Leu Pro
   1               5                  10                  15

 Pro Leu Leu Leu Leu Leu Leu Ala Ala Pro Trp Gly Arg Ala Val
                20                  25                  30

 Pro Cys Val Ser Gly Gly Leu Pro Lys Pro Ala Asn Ile Thr Phe
                35                  40                  45

 Leu Ser Ile Asn Met Lys Asn Val Leu Gln Trp Thr Pro Pro Glu
                50                  55                  60

 Gly Leu Gln Gly Val Lys Val Thr Tyr Thr Val Gln Tyr Phe Ile
                65                  70                  75

 Tyr Gly Gln Lys Lys Trp Leu Asn Lys Ser Glu Cys Arg Asn Ile
                80                  85                  90

 Asn Arg Thr Tyr Cys Asp Leu Ser Ala Glu Thr Ser Asp Tyr Glu
                95                 100                 105

 His Gln Tyr Tyr Ala Lys Val Lys Ala Ile Trp Gly Thr Lys Cys
                110                 115                 120

 Ser Lys Trp Ala Glu Ser Gly Arg Phe Tyr Pro Phe Leu Glu Thr
                125                 130                 135
```

```
Gln Ile Gly Pro Pro Glu Val Ala Leu Thr Thr Asp Glu Lys Ser
                140                 145                 150

Ile Ser Val Val Leu Thr Ala Pro Glu Lys Trp Lys Arg Asn Pro
                155                 160                 165

Glu Asp Leu Pro Val Ser Met Gln Gln Ile Tyr Ser Asn Leu Lys
                170                 175                 180

Tyr Asn Val Ser Val Leu Asn Thr Lys Ser Asn Arg Thr Trp Ser
                185                 190                 195

Gln Cys Val Thr Asn His Thr Leu Val Leu Thr Trp Leu Glu Pro
                200                 205                 210

Asn Thr Leu Tyr Cys Val His Val Glu Ser Phe Val Pro Gly Pro
                215                 220                 225

Pro Arg Arg Ala Gln Pro Ser Glu Lys Gln Cys Ala Arg Thr Leu
                230                 235                 240

Lys Asp Gln Ser Ser Glu Phe Lys Ala Lys Ile Ile Phe Trp Tyr
                245                 250                 255

Val Leu Pro Ile Ser Ile Thr Val Phe Leu Phe Ser Val Met Gly
                260                 265                 270

Tyr Ser Ile Tyr Arg Tyr Ile His Val Gly Lys Glu Lys His Pro
                275                 280                 285

Ala Asn Leu Ile Leu Ile Tyr Gly Asn Glu Phe Asp Lys Arg Phe
                290                 295                 300

Phe Val Pro Ala Glu Lys Ile Val Ile Asn Phe Ile Thr Leu Asn
                305                 310                 315

Ile Ser Asp Asp Ser Lys Ile Ser His Gln Asp Met Ser Leu Leu
                320                 325                 330

Gly Lys Ser Ser Asp Val Ser Ser Leu Asn Asp Pro Gln Pro Ser
                335                 340                 345

Gly Asn Leu Arg Pro Pro Gln Glu Glu Glu Glu Val Lys His Leu
                350                 355                 360

Gly Tyr Ala Ser His Leu Met Glu Ile Phe Cys Asp Ser Glu Glu
                365                 370                 375

Asn Thr Glu Gly Thr Ser Leu Thr Gln Gln Glu Ser Leu Ser Arg
                380                 385                 390

Thr Ile Pro Pro Asp Lys Thr Val Ile Glu Tyr Glu Tyr Asp Val
                395                 400                 405

Arg Thr Thr Asp Ile Cys Ala Gly Pro Glu Glu Gln Glu Leu Ser
                410                 415                 420

Leu Gln Glu Glu Val Ser Thr Gln Gly Thr Leu Leu Glu Ser Gln
                425                 430                 435
```

```
 Ala Ala Leu Ala Val Leu Gly Pro Gln Thr Leu Gln Tyr Ser Tyr
                 440                 445                 450

 Thr Pro Gln Leu Gln Asp Leu Asp Pro Leu Ala Gln Glu His Thr
                 455                 460                 465

 Asp Ser Glu Glu Gly Pro Glu Glu Glu Pro Ser Thr Thr Leu Val
                 470                 475                 480

 Asp Trp Asp Pro Gln Thr Gly Arg Leu Cys Ile Pro Ser Leu Ser
                 485                 490                 495

 Ser Phe Asp Gln Asp Ser Glu Gly Cys Glu Pro Ser Glu Gly Asp
                 500                 505                 510

 Gly Leu Gly Glu Glu Gly Leu Leu Ser Arg Leu Tyr Glu Glu Pro
                 515                 520                 525

 Ala Pro Asp Arg Pro Pro Gly Glu Asn Glu Thr Tyr Leu Met Gln
                 530                 535                 540

 Phe Met Glu Glu Trp Gly Leu Tyr Val Gln Met Glu Asn
                 545                 550

<210> 86
<211> 95
<212> PRT
<213> Homo sapiens

<400> 86
 Met Ala Ala Ala Arg Leu Cys Leu Ser Leu Leu Leu Leu Ser Thr
   1               5                  10                  15

 Cys Val Ala Leu Leu Leu Gln Pro Leu Leu Gly Ala Gln Gly Ala
                  20                  25                  30

 Pro Leu Glu Pro Val Tyr Pro Gly Asp Asn Ala Thr Pro Glu Gln
                  35                  40                  45

 Met Ala Gln Tyr Ala Ala Asp Leu Arg Arg Tyr Ile Asn Met Leu
                  50                  55                  60

 Thr Arg Pro Arg Tyr Gly Lys Arg His Lys Glu Asp Thr Leu Ala
                  65                  70                  75

 Phe Ser Glu Trp Gly Ser Pro His Ala Ala Val Pro Arg Glu Leu
                  80                  85                  90

 Ser Pro Leu Asp Leu
                  95

<210> 87
<211> 344
<212> PRT
<213> Homo sapiens

<400> 87
 Met Gly Pro Pro Ser Ala Pro Pro Cys Arg Leu His Val Pro Trp
   1               5                  10                  15

 Lys Glu Val Leu Leu Thr Ala Ser Leu Leu Thr Phe Trp Asn Pro
```

```
                20                  25                  30

Pro Thr Thr Ala Lys Leu Thr Ile Glu Ser Thr Pro Phe Asn Val
                35                  40                  45

Ala Glu Gly Lys Glu Val Leu Leu Leu Ala His Asn Leu Pro Gln
                50                  55                  60

Asn Arg Ile Gly Tyr Ser Trp Tyr Lys Gly Glu Arg Val Asp Gly
                65                  70                  75

Asn Ser Leu Ile Val Gly Tyr Val Ile Gly Thr Gln Gln Ala Thr
                80                  85                  90

Pro Gly Pro Ala Tyr Ser Gly Arg Glu Thr Ile Tyr Pro Asn Ala
                95                 100                 105

Ser Leu Leu Ile Gln Asn Val Thr Gln Asn Asp Thr Gly Phe Tyr
               110                 115                 120

Thr Leu Gln Val Ile Lys Ser Asp Leu Val Asn Glu Glu Ala Thr
               125                 130                 135

Gly Gln Phe His Val Tyr Pro Glu Leu Pro Lys Pro Ser Ile Ser
               140                 145                 150

Ser Asn Asn Ser Asn Pro Val Glu Asp Lys Asp Ala Val Ala Phe
               155                 160                 165

Thr Cys Glu Pro Glu Val Gln Asn Thr Thr Tyr Leu Trp Trp Val
               170                 175                 180

Asn Gly Gln Ser Leu Pro Val Ser Pro Arg Leu Gln Leu Ser Asn
               185                 190                 195

Gly Asn Met Thr Leu Thr Leu Leu Ser Val Lys Arg Asn Asp Ala
               200                 205                 210

Gly Ser Tyr Glu Cys Glu Ile Gln Asn Pro Ala Ser Ala Asn Arg
               215                 220                 225

Ser Asp Pro Val Thr Leu Asn Val Leu Tyr Gly Pro Asp Gly Pro
               230                 235                 240

Thr Ile Ser Pro Ser Lys Ala Asn Tyr Arg Pro Gly Glu Asn Leu
               245                 250                 255

Asn Leu Ser Cys His Ala Ala Ser Asn Pro Pro Ala Gln Tyr Ser
               260                 265                 270

Trp Phe Ile Asn Gly Thr Phe Gln Gln Ser Thr Gln Glu Leu Phe
               275                 280                 285

Ile Pro Asn Ile Thr Val Asn Asn Ser Gly Ser Tyr Met Cys Gln
               290                 295                 300

Ala His Asn Ser Ala Thr Gly Leu Asn Arg Thr Thr Val Thr Met
               305                 310                 315

Ile Thr Val Ser Gly Ser Ala Pro Val Leu Ser Ala Val Ala Thr
               320                 325                 330
```

```
 Val Gly Ile Thr Ile Gly Val Leu Ala Arg Val Ala Leu Ile
                 335                 340

<210> 88
<211> 987
<212> PRT
<213> Homo sapiens

<400> 88
 Met Glu Leu Arg Val Leu Leu Cys Trp Ala Ser Leu Ala Ala Ala
   1               5                  10                  15

 Leu Glu Glu Thr Leu Leu Asn Thr Lys Leu Glu Thr Ala Asp Leu
                  20                  25                  30

 Lys Trp Val Thr Phe Pro Gln Val Asp Gly Gln Trp Glu Glu Leu
                  35                  40                  45

 Ser Gly Leu Asp Glu Glu Gln His Ser Val Arg Thr Tyr Glu Val
                  50                  55                  60

 Cys Asp Val Gln Arg Ala Pro Gly Gln Ala His Trp Leu Arg Thr
                  65                  70                  75

 Gly Trp Val Pro Arg Arg Gly Ala Val His Val Tyr Ala Thr Leu
                  80                  85                  90

 Arg Phe Thr Met Leu Glu Cys Leu Ser Leu Pro Arg Ala Gly Arg
                  95                 100                 105

 Ser Cys Lys Glu Thr Phe Thr Val Phe Tyr Tyr Glu Ser Asp Ala
                 110                 115                 120

 Asp Thr Ala Thr Ala Leu Thr Pro Ala Trp Met Glu Asn Pro Tyr
                 125                 130                 135

 Ile Lys Val Asp Thr Val Ala Ala Glu His Leu Thr Arg Lys Arg
                 140                 145                 150

 Pro Gly Ala Glu Ala Thr Gly Lys Val Asn Val Lys Thr Leu Arg
                 155                 160                 165

 Leu Gly Pro Leu Ser Lys Ala Gly Phe Tyr Leu Ala Phe Gln Asp
                 170                 175                 180

 Gln Gly Ala Cys Met Ala Leu Leu Ser Leu His Leu Phe Tyr Lys
                 185                 190                 195

 Lys Cys Ala Gln Leu Thr Val Asn Leu Thr Arg Phe Pro Glu Thr
                 200                 205                 210

 Val Pro Arg Glu Leu Val Val Pro Val Ala Gly Ser Cys Val Val
                 215                 220                 225

 Asp Ala Val Pro Ala Pro Gly Pro Ser Pro Ser Leu Tyr Cys Arg
                 230                 235                 240

 Glu Asp Gly Gln Trp Ala Glu Gln Pro Val Thr Gly Cys Ser Cys
                 245                 250                 255
```

```
Ala Pro Gly Phe Glu Ala Ala Glu Gly Asn Thr Lys Cys Arg Ala
                260                 265                 270

Cys Ala Gln Gly Thr Phe Lys Pro Leu Ser Gly Glu Gly Ser Cys
                275                 280                 285

Gln Pro Cys Pro Ala Asn Ser His Ser Asn Thr Ile Gly Ser Ala
                290                 295                 300

Val Cys Gln Cys Arg Val Gly Tyr Phe Arg Ala Arg Thr Asp Pro
                305                 310                 315

Arg Gly Ala Pro Cys Thr Thr Pro Pro Ser Ala Pro Arg Ser Val
                320                 325                 330

Val Ser Arg Leu Asn Gly Ser Ser Leu His Leu Glu Trp Ser Ala
                335                 340                 345

Pro Leu Glu Ser Gly Gly Arg Glu Asp Leu Thr Tyr Ala Leu Arg
                350                 355                 360

Cys Arg Glu Cys Arg Pro Gly Gly Ser Cys Ala Pro Cys Gly Gly
                365                 370                 375

Asp Leu Thr Phe Asp Pro Gly Pro Arg Asp Leu Val Glu Pro Trp
                380                 385                 390

Val Val Val Arg Gly Leu Arg Pro Asp Phe Thr Tyr Thr Phe Glu
                395                 400                 405

Val Thr Ala Leu Asn Gly Val Ser Ser Leu Ala Thr Gly Pro Val
                410                 415                 420

Pro Phe Glu Pro Val Asn Val Thr Thr Asp Arg Glu Val Pro Pro
                425                 430                 435

Ala Val Ser Asp Ile Arg Val Thr Arg Ser Ser Pro Ser Ser Leu
                440                 445                 450

Ser Leu Ala Trp Ala Val Pro Arg Ala Pro Ser Gly Ala Val Leu
                455                 460                 465

Asp Tyr Glu Val Lys Tyr His Glu Lys Gly Ala Glu Gly Pro Ser
                470                 475                 480

Ser Val Arg Phe Leu Lys Thr Ser Glu Asn Arg Ala Glu Leu Arg
                485                 490                 495

Gly Leu Lys Arg Gly Ala Ser Tyr Leu Val Gln Val Arg Ala Arg
                500                 505                 510

Ser Glu Ala Gly Tyr Gly Pro Phe Gly Gln Glu His His Ser Gln
                515                 520                 525

Thr Gln Leu Asp Glu Ser Glu Gly Trp Arg Glu Gln Leu Ala Leu
                530                 535                 540

Ile Ala Gly Thr Ala Val Val Gly Val Val Leu Val Leu Val Val
                545                 550                 555

Ile Val Val Ala Val Leu Cys Leu Arg Lys Gln Ser Asn Gly Arg
```

```
                560                 565                 570

Glu Ala Glu Tyr Ser Asp Lys His Gly Gln Tyr Leu Ile Gly His
                575                 580                 585

Gly Thr Lys Val Tyr Ile Asp Pro Phe Thr Tyr Glu Asp Pro Asn
                590                 595                 600

Glu Ala Val Arg Glu Phe Ala Lys Glu Ile Asp Val Ser Tyr Val
                605                 610                 615

Lys Ile Glu Glu Val Ile Gly Ala Gly Glu Phe Gly Glu Val Cys
                620                 625                 630

Arg Gly Arg Leu Lys Ala Pro Gly Lys Lys Glu Ser Cys Val Ala
                635                 640                 645

Ile Lys Thr Leu Lys Gly Gly Tyr Thr Glu Arg Gln Arg Arg Glu
                650                 655                 660

Phe Leu Ser Glu Ala Ser Ile Met Gly Gln Phe Glu His Pro Asn
                665                 670                 675

Ile Ile Arg Leu Glu Gly Val Val Thr Asn Ser Met Pro Val Met
                680                 685                 690

Ile Leu Thr Glu Phe Met Glu Asn Gly Ala Leu Asp Ser Phe Leu
                695                 700                 705

Arg Leu Asn Asp Gly Gln Phe Thr Val Ile Gln Leu Val Gly Met
                710                 715                 720

Leu Arg Gly Ile Ala Ser Gly Met Arg Tyr Leu Ala Glu Met Ser
                725                 730                 735

Tyr Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Asn Ser
                740                 745                 750

Asn Leu Val Cys Lys Val Ser Asp Phe Gly Leu Ser Arg Phe Leu
                755                 760                 765

Glu Glu Asn Ser Ser Asp Pro Thr Tyr Thr Ser Ser Leu Gly Gly
                770                 775                 780

Lys Ile Pro Ile Arg Trp Thr Ala Pro Glu Ala Ile Ala Phe Arg
                785                 790                 795

Lys Phe Thr Ser Ala Ser Asp Ala Trp Ser Tyr Gly Ile Val Met
                800                 805                 810

Trp Glu Val Met Ser Phe Gly Glu Arg Pro Tyr Trp Asp Met Ser
                815                 820                 825

Asn Gln Asp Val Ile Asn Ala Ile Glu Gln Asp Tyr Arg Leu Pro
                830                 835                 840

Pro Pro Pro Asp Cys Pro Thr Ser Leu His Gln Leu Met Leu Asp
                845                 850                 855

Cys Trp Gln Lys Asp Arg Asn Ala Arg Pro Arg Phe Pro Gln Val
                860                 865                 870
```

```
 Val Ser Ala Leu Asp Lys Met Ile Arg Asn Pro Ala Ser Leu Lys
                 875                 880                 885

 Ile Val Ala Arg Glu Asn Gly Gly Ala Ser His Pro Leu Leu Asp
                 890                 895                 900

 Gln Arg Gln Pro His Tyr Ser Ala Phe Gly Ser Val Gly Glu Trp
                 905                 910                 915

 Leu Arg Ala Ile Lys Met Gly Arg Tyr Glu Glu Ser Phe Ala Ala
                 920                 925                 930

 Ala Gly Phe Gly Ser Phe Glu Leu Val Ser Gln Ile Ser Ala Glu
                 935                 940                 945

 Asp Leu Leu Arg Ile Gly Val Thr Leu Ala Gly His Gln Lys Lys
                 950                 955                 960

 Ile Leu Ala Ser Val Gln His Met Lys Ser Gln Ala Lys Pro Gly
                 965                 970                 975

 Thr Pro Gly Gly Thr Gly Gly Pro Ala Pro Gln Tyr
                 980                 985

<210> 89
<211> 911
<212> PRT
<213> Homo sapiens

<400> 89
 Met Ala Gln Leu Phe Leu Pro Leu Leu Ala Ala Leu Val Leu Ala
   1               5                  10                  15

 Gln Ala Pro Ala Ala Leu Ala Asp Val Leu Glu Gly Asp Ser Ser
                  20                  25                  30

 Glu Asp Arg Ala Phe Arg Val Arg Ile Ala Gly Asp Ala Pro Leu
                  35                  40                  45

 Gln Gly Val Leu Gly Gly Ala Leu Thr Ile Pro Cys His Val His
                  50                  55                  60

 Tyr Leu Arg Pro Pro Pro Ser Arg Arg Ala Val Leu Gly Ser Pro
                  65                  70                  75

 Arg Val Lys Trp Thr Phe Leu Ser Arg Gly Arg Glu Ala Glu Val
                  80                  85                  90

 Leu Val Ala Arg Gly Val Arg Val Lys Val Asn Glu Ala Tyr Arg
                  95                 100                 105

 Phe Arg Val Ala Leu Pro Ala Tyr Pro Ala Ser Leu Thr Asp Val
                 110                 115                 120

 Ser Leu Ala Leu Ser Glu Leu Arg Pro Asn Asp Ser Gly Ile Tyr
                 125                 130                 135

 Arg Cys Glu Val Gln His Gly Ile Asp Asp Ser Ser Asp Ala Val
                 140                 145                 150
```

```
Glu Val Lys Val Lys Gly Val Val Phe Leu Tyr Arg Glu Gly Ser
                155                 160                 165

Ala Arg Tyr Ala Phe Ser Phe Ser Gly Ala Gln Glu Ala Cys Ala
                170                 175                 180

Arg Ile Gly Ala His Ile Ala Thr Pro Glu Gln Leu Tyr Ala Ala
                185                 190                 195

Tyr Leu Gly Gly Tyr Glu Gln Cys Asp Ala Gly Trp Leu Ser Asp
                200                 205                 210

Gln Thr Val Arg Tyr Pro Ile Gln Thr Pro Arg Glu Ala Cys Tyr
                215                 220                 225

Gly Asp Met Asp Gly Phe Pro Gly Val Arg Asn Tyr Gly Val Val
                230                 235                 240

Asp Pro Asp Asp Leu Tyr Asp Val Tyr Cys Tyr Ala Glu Asp Leu
                245                 250                 255

Asn Gly Glu Leu Phe Leu Gly Asp Pro Pro Glu Lys Leu Thr Leu
                260                 265                 270

Glu Glu Ala Arg Ala Tyr Cys Gln Glu Arg Gly Ala Glu Ile Ala
                275                 280                 285

Thr Thr Gly Gln Leu Tyr Ala Ala Trp Asp Gly Gly Leu Asp His
                290                 295                 300

Cys Ser Pro Gly Trp Leu Ala Asp Gly Ser Val Arg Tyr Pro Ile
                305                 310                 315

Val Thr Pro Ser Gln Arg Cys Gly Gly Gly Leu Pro Gly Val Lys
                320                 325                 330

Thr Leu Phe Leu Phe Pro Asn Gln Thr Gly Phe Pro Asn Lys His
                335                 340                 345

Ser Arg Phe Asn Val Tyr Cys Phe Arg Asp Ser Ala Gln Pro Ser
                350                 355                 360

Ala Ile Pro Glu Ala Ser Asn Pro Ala Ser Asn Pro Ala Ser Asp
                365                 370                 375

Gly Leu Glu Ala Ile Val Thr Val Thr Glu Thr Leu Glu Glu Leu
                380                 385                 390

Gln Leu Pro Gln Glu Ala Thr Glu Ser Glu Ser Arg Gly Ala Ile
                395                 400                 405

Tyr Ser Ile Pro Ile Met Glu Asp Gly Gly Gly Gly Ser Ser Thr
                410                 415                 420

Pro Glu Asp Pro Ala Glu Ala Pro Arg Thr Leu Leu Glu Phe Glu
                425                 430                 435

Thr Gln Ser Met Val Pro Pro Thr Gly Phe Ser Glu Glu Glu Gly
                440                 445                 450

Lys Ala Leu Glu Glu Glu Glu Lys Tyr Glu Asp Glu Glu Glu Lys
```

```
                455                 460                 465

Glu Glu Glu Glu Glu Glu Glu Glu Val Glu Asp Glu Ala Leu Trp
                470                 475                 480

Ala Trp Pro Ser Glu Leu Ser Ser Pro Gly Pro Glu Ala Ser Leu
                485                 490                 495

Pro Thr Glu Pro Ala Ala Gln Glu Lys Ser Leu Ser Gln Ala Pro
                500                 505                 510

Ala Arg Ala Val Leu Gln Pro Gly Ala Ser Pro Leu Pro Asp Gly
                515                 520                 525

Glu Ser Glu Ala Ser Arg Pro Pro Arg Val His Gly Pro Pro Thr
                530                 535                 540

Glu Thr Leu Pro Thr Pro Arg Glu Arg Asn Leu Ala Ser Pro Ser
                545                 550                 555

Pro Ser Thr Leu Val Glu Ala Arg Glu Val Gly Glu Ala Thr Gly
                560                 565                 570

Gly Pro Glu Leu Ser Gly Val Pro Arg Gly Glu Ser Glu Glu Thr
                575                 580                 585

Gly Ser Ser Glu Gly Ala Pro Ser Leu Leu Pro Ala Thr Arg Ala
                590                 595                 600

Pro Glu Gly Thr Arg Glu Leu Glu Ala Pro Ser Glu Asp Asn Ser
                605                 610                 615

Gly Arg Thr Ala Pro Ala Gly Thr Ser Val Gln Ala Gln Pro Val
                620                 625                 630

Leu Pro Thr Asp Ser Ala Ser Arg Gly Gly Val Ala Val Val Pro
                635                 640                 645

Ala Ser Gly Asp Cys Val Pro Ser Pro Cys His Asn Gly Gly Thr
                650                 655                 660

Cys Leu Glu Glu Glu Glu Gly Val Arg Cys Leu Cys Leu Pro Gly
                665                 670                 675

Tyr Gly Gly Asp Leu Cys Asp Val Gly Leu Arg Phe Cys Asn Pro
                680                 685                 690

Gly Trp Asp Ala Phe Gln Gly Ala Cys Tyr Lys His Phe Ser Thr
                695                 700                 705

Arg Arg Ser Trp Glu Glu Ala Glu Thr Gln Cys Arg Met Tyr Gly
                710                 715                 720

Ala His Leu Ala Ser Ile Ser Thr Pro Glu Glu Gln Asp Phe Ile
                725                 730                 735

Asn Asn Arg Tyr Arg Glu Tyr Gln Trp Ile Gly Leu Asn Asp Arg
                740                 745                 750

Thr Ile Glu Gly Asp Phe Leu Trp Ser Asp Gly Val Pro Leu Leu
                755                 760                 765
```

```
 Tyr Glu Asn Trp Asn Pro Gly Gln Pro Asp Ser Tyr Phe Leu Ser
                 770                 775                 780

 Gly Glu Asn Cys Val Val Met Val Trp His Asp Gln Gly Gln Trp
                 785                 790                 795

 Ser Asp Val Pro Cys Asn Tyr His Leu Ser Tyr Thr Cys Lys Met
                 800                 805                 810

 Gly Leu Val Ser Cys Gly Pro Pro Pro Glu Leu Pro Leu Ala Gln
                 815                 820                 825

 Val Phe Gly Arg Pro Arg Leu Arg Tyr Glu Val Asp Thr Val Leu
                 830                 835                 840

 Arg Tyr Arg Cys Arg Glu Gly Leu Ala Gln Arg Asn Leu Pro Leu
                 845                 850                 855

 Ile Arg Cys Gln Glu Asn Gly Arg Trp Glu Ala Pro Gln Ile Ser
                 860                 865                 870

 Cys Val Pro Arg Arg Pro Ala Arg Ala Leu His Pro Glu Glu Asp
                 875                 880                 885

 Pro Glu Gly Arg Gln Gly Arg Leu Leu Gly Arg Trp Lys Ala Leu
                 890                 895                 900

 Leu Ile Pro Pro Ser Ser Pro Met Pro Gly Pro
                 905                 910

<210> 90
<211> 671
<212> PRT
<213> Homo sapiens

<400> 90
 Met Ala Gln Leu Phe Leu Pro Leu Leu Ala Ala Leu Val Leu Ala
   1               5                  10                  15

 Gln Ala Pro Ala Ala Leu Ala Asp Val Leu Glu Gly Asp Ser Ser
                  20                  25                  30

 Glu Asp Arg Ala Phe Arg Val Arg Ile Ala Gly Asp Ala Pro Leu
                  35                  40                  45

 Gln Gly Val Leu Gly Gly Ala Leu Thr Ile Pro Cys His Val His
                  50                  55                  60

 Tyr Leu Arg Pro Pro Pro Ser Arg Arg Ala Val Leu Gly Ser Pro
                  65                  70                  75

 Arg Val Lys Trp Thr Phe Leu Ser Arg Gly Arg Glu Ala Glu Val
                  80                  85                  90

 Leu Val Ala Arg Gly Val Arg Val Lys Val Asn Glu Ala Tyr Arg
                  95                 100                 105

 Phe Arg Val Ala Leu Pro Ala Tyr Pro Ala Ser Leu Thr Asp Val
                 110                 115                 120
```

```
Ser Leu Ala Leu Ser Glu Leu Arg Pro Asn Asp Ser Gly Ile Tyr
                125                 130                 135

Arg Cys Glu Val Gln His Gly Ile Asp Asp Ser Ser Asp Ala Val
                140                 145                 150

Glu Val Lys Val Lys Gly Val Val Phe Leu Tyr Arg Glu Gly Ser
                155                 160                 165

Ala Arg Tyr Ala Phe Ser Phe Ser Gly Ala Gln Glu Ala Cys Ala
                170                 175                 180

Arg Ile Gly Ala His Ile Ala Thr Pro Glu Gln Leu Tyr Ala Ala
                185                 190                 195

Tyr Leu Gly Gly Tyr Glu Gln Cys Asp Ala Gly Trp Leu Ser Asp
                200                 205                 210

Gln Thr Val Arg Tyr Pro Ile Gln Thr Pro Arg Glu Ala Cys Tyr
                215                 220                 225

Gly Asp Met Asp Gly Phe Pro Gly Val Arg Asn Tyr Gly Val Val
                230                 235                 240

Asp Pro Asp Asp Leu Tyr Asp Val Tyr Cys Tyr Ala Glu Asp Leu
                245                 250                 255

Asn Gly Glu Leu Phe Leu Gly Asp Pro Pro Glu Lys Leu Thr Leu
                260                 265                 270

Glu Glu Ala Arg Ala Tyr Cys Gln Glu Arg Gly Ala Glu Ile Ala
                275                 280                 285

Thr Thr Gly Gln Leu Tyr Ala Ala Trp Asp Gly Gly Leu Asp His
                290                 295                 300

Cys Ser Pro Gly Trp Leu Ala Asp Gly Ser Val Arg Tyr Pro Ile
                305                 310                 315

Val Thr Pro Ser Gln Arg Cys Gly Gly Gly Leu Pro Gly Val Lys
                320                 325                 330

Thr Leu Phe Leu Phe Pro Asn Gln Thr Gly Phe Pro Asn Lys His
                335                 340                 345

Ser Arg Phe Asn Val Tyr Cys Phe Arg Asp Ser Ala Gln Pro Ser
                350                 355                 360

Ala Ile Pro Glu Ala Ser Asn Pro Ala Ser Asn Pro Ala Ser Asp
                365                 370                 375

Gly Leu Glu Ala Ile Val Thr Val Thr Glu Thr Leu Glu Glu Leu
                380                 385                 390

Gln Leu Pro Gln Glu Ala Thr Glu Ser Glu Ser Arg Gly Ala Ile
                395                 400                 405

Tyr Ser Ile Pro Ile Met Glu Asp Gly Gly Gly Gly Ser Ser Thr
                410                 415                 420

Pro Glu Asp Pro Ala Glu Ala Pro Arg Thr Leu Leu Glu Phe Glu
```

```
                425                 430                 435

 Thr Gln Ser Met Val Pro Pro Thr Gly Phe Ser Glu Glu Glu Gly
                440                 445                 450

 Lys Ala Leu Glu Glu Glu Glu Lys Tyr Glu Asp Glu Glu Glu Lys
                455                 460                 465

 Glu Glu Glu Glu Glu Glu Glu Glu Val Glu Asp Glu Ala Leu Trp
                470                 475                 480

 Ala Trp Pro Ser Glu Leu Ser Ser Pro Gly Pro Glu Ala Ser Leu
                485                 490                 495

 Pro Thr Glu Pro Ala Ala Gln Glu Glu Ser Leu Ser Gln Ala Pro
                500                 505                 510

 Ala Arg Ala Val Leu Gln Pro Gly Ala Ser Pro Leu Pro Asp Gly
                515                 520                 525

 Glu Ser Glu Ala Ser Arg Pro Pro Arg Val His Gly Pro Pro Thr
                530                 535                 540

 Glu Thr Leu Pro Thr Pro Arg Glu Arg Asn Leu Ala Ser Pro Ser
                545                 550                 555

 Pro Ser Thr Leu Val Glu Ala Arg Glu Val Gly Glu Ala Thr Gly
                560                 565                 570

 Gly Pro Glu Leu Ser Gly Val Pro Arg Gly Glu Ser Glu Glu Thr
                575                 580                 585

 Gly Ser Ser Glu Gly Ala Pro Ser Leu Leu Pro Ala Thr Arg Ala
                590                 595                 600

 Pro Glu Gly Thr Arg Glu Leu Glu Ala Pro Ser Glu Asp Asn Ser
                605                 610                 615

 Gly Arg Thr Ala Pro Ala Gly Thr Ser Val Gln Ala Gln Pro Val
                620                 625                 630

 Leu Pro Thr Asp Ser Ala Ser Arg Gly Gly Val Ala Val Val Pro
                635                 640                 645

 Ala Ser Gly Asn Ser Ala Gln Gly Ser Thr Ala Leu Ser Ile Leu
                650                 655                 660

 Leu Leu Phe Phe Pro Leu Gln Leu Trp Val Thr
                665                 670

<210> 91
<211> 86
<212> PRT
<213> Homo sapiens

<400> 91
 Met Ala Val Arg Gln Trp Val Ile Ala Leu Ala Leu Ala Ala Leu
   1               5                  10                  15

 Leu Val Val Asp Arg Glu Val Pro Val Ala Ala Gly Lys Leu Pro
                 20                  25                  30
```

```
 Phe Ser Arg Met Pro Ile Cys Glu His Met Val Glu Ser Pro Thr
                  35                  40                  45

 Cys Ser Gln Met Ser Asn Leu Val Cys Gly Thr Asp Gly Leu Thr
                  50                  55                  60

 Tyr Thr Asn Glu Cys Gln Leu Cys Leu Ala Arg Ile Lys Thr Lys
                  65                  70                  75

 Gln Asp Ile Gln Ile Met Lys Asp Gly Lys Cys
                  80                  85

<210> 92
<211> 459
<212> PRT
<213> Homo sapiens

<400> 92
 Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu Ile Pro
   1               5                  10                  15

 Ala Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu Leu
                  20                  25                  30

 Leu Leu Val Pro Val His Pro Gln Arg Leu Pro Arg Met Gln Glu
                  35                  40                  45

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

 Gly Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu
                  65                  70                  75

 Asp Pro Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro
                  80                  85                  90

 Gly Glu Glu Asp Leu Pro Glu Val Lys Pro Lys Ser Glu Glu Glu
                  95                 100                 105

 Gly Ser Leu Lys Leu Glu Asp Leu Pro Thr Val Glu Ala Pro Gly
                 110                 115                 120

 Asp Pro Gln Glu Pro Gln Asn Asn Ala His Arg Asp Lys Glu Gly
                 125                 130                 135

 Asp Asp Gln Ser His Trp Arg Tyr Gly Gly Asp Pro Pro Trp Pro
                 140                 145                 150

 Arg Val Ser Pro Ala Cys Ala Gly Arg Phe Gln Ser Pro Val Asp
                 155                 160                 165

 Ile Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala Leu Arg Pro Leu
                 170                 175                 180

 Glu Leu Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu Leu Arg Leu
                 185                 190                 195

 Arg Asn Asn Gly His Ser Val Gln Leu Thr Leu Pro Pro Gly Leu
                 200                 205                 210
```

```
 Glu Met Ala Leu Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln Leu
                 215                 220                 225

 His Leu His Trp Gly Ala Ala Gly Arg Pro Gly Ser Glu His Thr
                 230                 235                 240

 Val Glu Gly His Arg Phe Pro Ala Glu Ile His Val Val His Leu
                 245                 250                 255

 Ser Thr Ala Phe Ala Arg Val Asp Glu Ala Leu Gly Arg Pro Gly
                 260                 265                 270

 Gly Leu Ala Val Leu Ala Ala Phe Leu Glu Glu Gly Pro Glu Glu
                 275                 280                 285

 Asn Ser Ala Tyr Glu Gln Leu Leu Ser Arg Leu Glu Glu Ile Ala
                 290                 295                 300

 Glu Glu Gly Ser Glu Thr Gln Val Pro Gly Leu Asp Ile Ser Ala
                 305                 310                 315

 Leu Leu Pro Ser Asp Phe Ser Arg Tyr Phe Arg Tyr Glu Gly Ser
                 320                 325                 330

 Leu Thr Thr Pro Pro Cys Ala Gln Gly Val Ile Trp Thr Val Phe
                 335                 340                 345

 Asn Gln Thr Val Met Leu Ser Ala Lys Gln Leu His Thr Leu Ser
                 350                 355                 360

 Asp Thr Leu Trp Gly Pro Gly Asp Ser Arg Leu Gln Leu Asn Phe
                 365                 370                 375

 Arg Ala Thr Gln Pro Leu Asn Gly Arg Val Ile Glu Ala Ser Phe
                 380                 385                 390

 Pro Ala Gly Val Asp Ser Ser Pro Arg Ala Ala Glu Pro Val Gln
                 395                 400                 405

 Leu Asn Ser Cys Leu Ala Ala Gly Asp Ile Leu Ala Leu Val Phe
                 410                 415                 420

 Gly Leu Leu Phe Ala Val Thr Ser Val Ala Phe Leu Val Gln Met
                 425                 430                 435

 Arg Arg Gln His Arg Arg Gly Thr Lys Gly Gly Val Ser Tyr Arg
                 440                 445                 450

 Pro Ala Glu Val Ala Glu Thr Gly Ala
                 455

<210> 93
<211> 375
<212> PRT
<213> Homo sapiens

<400> 93
 Met Asp Ala Leu Gln Leu Ala Asn Ser Ala Phe Ala Val Asp Leu
   1               5                  10                  15

 Phe Lys Gln Leu Cys Glu Lys Glu Pro Leu Gly Asn Val Leu Phe
```

```
                20                  25                  30

Ser Pro Ile Cys Leu Ser Thr Ser Leu Ser Leu Ala Gln Val Gly
                35                  40                  45

Ala Lys Gly Asp Thr Ala Asn Glu Ile Gly Gln Val Leu His Phe
                50                  55                  60

Glu Asn Val Lys Asp Val Pro Phe Gly Phe Gln Thr Val Thr Ser
                65                  70                  75

Asp Val Asn Lys Leu Ser Ser Phe Tyr Ser Leu Lys Leu Ile Lys
                80                  85                  90

Arg Leu Tyr Val Asp Lys Ser Leu Asn Leu Ser Thr Glu Phe Ile
                95                 100                 105

Ser Ser Thr Lys Arg Pro Tyr Ala Lys Glu Leu Glu Thr Val Asp
               110                 115                 120

Phe Lys Asp Lys Leu Glu Glu Thr Lys Gly Gln Ile Asn Asn Ser
               125                 130                 135

Ile Lys Asp Leu Thr Asp Gly His Phe Glu Asn Ile Leu Ala Asp
               140                 145                 150

Asn Ser Val Asn Asp Gln Thr Lys Ile Leu Val Val Asn Ala Ala
               155                 160                 165

Tyr Phe Val Gly Lys Trp Met Lys Lys Phe Ser Glu Ser Glu Thr
               170                 175                 180

Lys Glu Cys Pro Phe Arg Val Asn Lys Thr Asp Thr Lys Pro Val
               185                 190                 195

Gln Met Met Asn Met Glu Ala Thr Phe Cys Met Gly Asn Ile Asp
               200                 205                 210

Ser Ile Asn Cys Lys Ile Ile Glu Leu Pro Phe Gln Asn Lys His
               215                 220                 225

Leu Ser Met Phe Ile Leu Leu Pro Lys Asp Val Glu Asp Glu Ser
               230                 235                 240

Thr Gly Leu Glu Lys Ile Glu Lys Gln Leu Asn Ser Glu Ser Leu
               245                 250                 255

Ser Gln Trp Thr Asn Pro Ser Thr Met Ala Asn Ala Lys Val Lys
               260                 265                 270

Leu Ser Ile Pro Lys Phe Lys Val Glu Lys Met Ile Asp Pro Lys
               275                 280                 285

Ala Cys Leu Glu Asn Leu Gly Leu Lys His Ile Phe Ser Glu Asp
               290                 295                 300

Thr Ser Asp Phe Ser Gly Met Ser Glu Thr Lys Gly Val Ala Leu
               305                 310                 315

Ser Asn Val Ile His Lys Val Cys Leu Glu Ile Thr Glu Asp Gly
               320                 325                 330
```

```
 Gly Asp Ser Ile Glu Val Pro Gly Ala Arg Ile Leu Gln His Lys
                 335                 340                 345

 Asp Glu Leu Asn Ala Asp His Pro Phe Ile Tyr Ile Ile Arg His
                 350                 355                 360

 Asn Lys Thr Arg Asn Ile Ile Phe Phe Gly Lys Phe Cys Ser Pro
                 365                 370                 375

<210> 94
<211> 260
<212> PRT
<213> Homo sapiens

<400> 94
 Met Asp Arg His Ser Ser Tyr Ile Phe Ile Trp Leu Gln Leu Glu
   1               5                  10                  15

 Leu Cys Ala Met Ala Val Leu Leu Thr Lys Gly Glu Ile Arg Cys
                  20                  25                  30

 Tyr Cys Asp Ala Ala His Cys Val Ala Thr Gly Tyr Met Cys Lys
                  35                  40                  45

 Ser Glu Leu Ser Ala Cys Phe Ser Arg Leu Leu Asp Pro Gln Asn
                  50                  55                  60

 Ser Asn Ser Pro Leu Thr His Gly Cys Leu Asp Ser Leu Ala Ser
                  65                  70                  75

 Thr Thr Asp Ile Cys Gln Ala Lys Gln Ala Arg Asn His Ser Gly
                  80                  85                  90

 Thr Thr Ile Pro Thr Leu Glu Cys Cys His Glu Asp Met Cys Asn
                  95                 100                 105

 Tyr Arg Gly Leu His Asp Val Leu Ser Pro Pro Arg Gly Glu Ala
                 110                 115                 120

 Ser Gly Gln Gly Asn Arg Tyr Gln His Asp Gly Ser Arg Asn Leu
                 125                 130                 135

 Ile Thr Lys Val Gln Glu Leu Thr Ser Ser Lys Glu Leu Trp Phe
                 140                 145                 150

 Arg Ala Ala Val Ile Ala Val Pro Ile Ala Gly Gly Leu Ile Leu
                 155                 160                 165

 Val Leu Leu Ile Met Leu Ala Leu Arg Met Leu Arg Ser Glu Asn
                 170                 175                 180

 Lys Arg Leu Gln Asp Gln Arg Gln Gln Met Leu Ser Arg Leu His
                 185                 190                 195

 Tyr Ser Phe His Gly His His Ser Lys Lys Gly Gln Val Ala Lys
                 200                 205                 210

 Leu Asp Leu Glu Cys Met Val Pro Val Ser Gly His Glu Asn Cys
                 215                 220                 225
```

```
 Cys Leu Thr Cys Asp Lys Met Arg Gln Ala Asp Leu Ser Asn Asp
                 230                 235                 240

 Lys Ile Leu Ser Leu Val His Trp Gly Met Tyr Ser Gly His Gly
                 245                 250                 255

 Lys Leu Glu Phe Val
                 260
```

<210> 95
<211> 1095
<212> PRT
<213> Homo sapiens

<400> 95

```
 Met Arg Asn Arg Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu
   1               5                  10                  15

 Tyr Ser Ser Ala Ser Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser
                  20                  25                  30

 Asp Leu Val Asn Phe Ile Gln Ala Asn Phe Lys Lys Arg Glu Cys
                  35                  40                  45

 Val Phe Phe Thr Lys Asp Ser Lys Ala Thr Glu Asn Val Cys Lys
                  50                  55                  60

 Cys Gly Tyr Ala Gln Ser Gln His Met Glu Gly Thr Gln Ile Asn
                  65                  70                  75

 Gln Ser Glu Lys Trp Asn Tyr Lys Lys His Thr Lys Glu Phe Pro
                  80                  85                  90

 Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu Thr Leu Gly Lys Lys
                  95                 100                 105

 Gly Lys Tyr Ile Arg Leu Ser Cys Asp Thr Asp Ala Glu Ile Leu
                 110                 115                 120

 Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr Pro Asn Leu
                 125                 130                 135

 Val Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Ala Leu Lys Pro
                 140                 145                 150

 Arg Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln Ser
                 155                 160                 165

 Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His Tyr Gly Leu Thr
                 170                 175                 180

 Lys Tyr Ile Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser
                 185                 190                 195

 Ser Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Ala Trp Gly Met
                 200                 205                 210

 Val Ser Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly
                 215                 220                 225

 Tyr Phe Leu Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro
```

```
                230                 235                 240

Leu Tyr Ile Leu Asp Asn Asn His Thr His Leu Leu Leu Val Asp
                245                 250                 255

Asn Gly Cys His Gly His Pro Thr Val Glu Ala Lys Leu Arg Asn
                260                 265                 270

Gln Leu Glu Lys His Ile Ser Glu Arg Thr Ile Gln Asp Ser Asn
                275                 280                 285

Tyr Gly Gly Lys Ile Pro Ile Val Cys Phe Ala Gln Gly Gly Gly
                290                 295                 300

Lys Glu Thr Leu Lys Ala Ile Asn Thr Ser Ile Lys Asn Lys Ile
                305                 310                 315

Pro Cys Val Val Val Glu Gly Ser Gly Arg Ile Ala Asp Val Ile
                320                 325                 330

Ala Ser Leu Val Glu Val Glu Asp Ala Pro Thr Ser Ser Ala Val
                335                 340                 345

Lys Glu Lys Leu Val Arg Phe Leu Pro Arg Thr Val Ser Arg Leu
                350                 355                 360

Ser Glu Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu Lys Glu Ile
                365                 370                 375

Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu Glu Ala
                380                 385                 390

Gly Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr Lys
                395                 400                 405

Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln
                410                 415                 420

Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp
                425                 430                 435

Glu Ile Phe Thr Asn Asp Arg Arg Trp Glu Ser Ala Asp Leu Gln
                440                 445                 450

Glu Val Met Phe Thr Ala Leu Ile Lys Asp Arg Pro Lys Phe Val
                455                 460                 465

Arg Leu Phe Leu Glu Asn Gly Leu Asn Leu Arg Lys Phe Leu Thr
                470                 475                 480

His Asp Val Leu Thr Glu Leu Phe Ser Asn His Phe Ser Thr Leu
                485                 490                 495

Val Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser Tyr Asn Asp Ala
                500                 505                 510

Leu Leu Thr Phe Val Trp Lys Leu Val Ala Asn Phe Arg Arg Gly
                515                 520                 525

Phe Arg Lys Glu Asp Arg Asn Gly Arg Asp Glu Met Asp Ile Glu
                530                 535                 540
```

```
Leu His Asp Val Ser Pro Ile Thr Arg His Pro Leu Gln Ala Leu
                545                 550                 555

Phe Ile Trp Ala Ile Leu Gln Asn Lys Lys Glu Leu Ser Lys Val
                560                 565                 570

Ile Trp Glu Gln Thr Arg Gly Cys Thr Leu Ala Ala Leu Gly Ala
                575                 580                 585

Ser Lys Leu Leu Lys Thr Leu Ala Lys Val Lys Asn Asp Ile Asn
                590                 595                 600

Ala Ala Gly Glu Ser Glu Glu Leu Ala Asn Glu Tyr Glu Thr Arg
                605                 610                 615

Ala Val Glu Leu Phe Thr Glu Cys Tyr Ser Ser Asp Glu Asp Leu
                620                 625                 630

Ala Glu Gln Leu Leu Val Tyr Ser Cys Glu Ala Trp Gly Gly Ser
                635                 640                 645

Asn Cys Leu Glu Leu Ala Val Glu Ala Thr Asp Gln His Phe Thr
                650                 655                 660

Ala Gln Pro Gly Val Gln Asn Phe Leu Ser Lys Gln Trp Tyr Gly
                665                 670                 675

Glu Ile Ser Arg Asp Thr Lys Asn Trp Lys Ile Ile Leu Cys Leu
                680                 685                 690

Phe Ile Ile Pro Leu Val Gly Cys Gly Phe Val Ser Phe Arg Lys
                695                 700                 705

Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp Tyr Tyr Val Ala
                710                 715                 720

Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn Val Val Phe
                725                 730                 735

Tyr Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu Met Asp
                740                 745                 750

Phe His Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser Leu
                755                 760                 765

Val Phe Val Leu Phe Cys Asp Glu Val Arg Gln Trp Tyr Val Asn
                770                 775                 780

Gly Val Asn Tyr Phe Thr Asp Leu Trp Asn Val Met Asp Thr Leu
                785                 790                 795

Gly Leu Phe Tyr Phe Ile Ala Gly Ile Val Phe Arg Leu His Ser
                800                 805                 810

Ser Asn Lys Ser Ser Leu Tyr Ser Gly Arg Val Ile Phe Cys Leu
                815                 820                 825

Asp Tyr Ile Ile Phe Thr Leu Arg Leu Ile His Ile Phe Thr Val
                830                 835                 840
```

```
 Ser Arg Asn Leu Gly Pro Lys Ile Ile Met Leu Gln Arg Met Leu
                 845                 850                 855

 Ile Asp Val Phe Phe Phe Leu Phe Leu Phe Ala Val Trp Met Val
                 860                 865                 870

 Ala Phe Gly Val Ala Arg Gln Gly Ile Leu Arg Gln Asn Glu Gln
                 875                 880                 885

 Arg Trp Arg Trp Ile Phe Arg Ser Val Ile Tyr Glu Pro Tyr Leu
                 890                 895                 900

 Ala Met Phe Gly Gln Val Pro Ser Asp Val Asp Gly Thr Thr Tyr
                 905                 910                 915

 Asp Phe Ala His Cys Thr Phe Thr Gly Asn Glu Ser Lys Pro Leu
                 920                 925                 930

 Cys Val Glu Leu Asp Glu His Asn Leu Pro Arg Phe Pro Glu Trp
                 935                 940                 945

 Ile Thr Ile Pro Leu Val Cys Ile Tyr Met Leu Ser Thr Asn Ile
                 950                 955                 960

 Leu Leu Val Asn Leu Leu Val Ala Met Phe Gly Tyr Thr Val Gly
                 965                 970                 975

 Thr Val Gln Glu Asn Asn Asp Gln Val Trp Lys Phe Gln Arg Tyr
                 980                 985                 990

 Phe Leu Val Gln Glu Tyr Cys Ser Arg Leu Asn Ile Pro Phe Pro
                 995                1000                1005

 Phe Ile Val Phe Ala Tyr Phe Tyr Met Val Val Lys Lys Cys Phe
                1010                1015                1020

 Lys Cys Cys Cys Lys Glu Lys Asn Met Glu Ser Ser Val Cys Cys
                1025                1030                1035

 Phe Lys Asn Glu Asp Asn Glu Thr Leu Ala Trp Glu Gly Val Met
                1040                1045                1050

 Lys Glu Asn Tyr Leu Val Lys Ile Asn Thr Lys Ala Asn Asp Thr
                1055                1060                1065

 Ser Glu Glu Met Arg His Arg Phe Arg Gln Leu Asp Thr Lys Leu
                1070                1075                1080

 Asn Asp Leu Lys Gly Leu Leu Lys Glu Ile Ala Asn Lys Ile Lys
                1085                1090                1095

<210> 96
<211> 1095
<212> PRT
<213> Homo sapiens

<400> 96
 Met Arg Asn Arg Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu
   1               5                  10                  15

 Tyr Ser Ser Ala Ser Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser
```

```
                 20                  25                  30

Asp Leu Val Asn Phe Ile Gln Ala Asn Phe Lys Lys Arg Glu Cys
                 35                  40                  45

Val Phe Phe Thr Lys Asp Ser Lys Ala Thr Glu Asn Val Cys Lys
                 50                  55                  60

Cys Gly Tyr Ala Gln Ser Gln His Met Glu Gly Thr Gln Ile Asn
                 65                  70                  75

Gln Ser Glu Lys Trp Asn Tyr Lys Lys His Thr Lys Glu Phe Pro
                 80                  85                  90

Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu Thr Leu Gly Lys Lys
                 95                 100                 105

Gly Lys Tyr Ile Arg Leu Ser Cys Asp Thr Asp Ala Glu Ile Leu
                110                 115                 120

Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr Pro Asn Leu
                125                 130                 135

Val Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Ala Leu Lys Pro
                140                 145                 150

Arg Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln Ser
                155                 160                 165

Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His Tyr Gly Leu Met
                170                 175                 180

Lys Tyr Ile Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser
                185                 190                 195

Ser Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Ala Trp Gly Met
                200                 205                 210

Val Ser Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly
                215                 220                 225

Tyr Phe Leu Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro
                230                 235                 240

Leu Tyr Ile Leu Asp Asn Asn His Thr His Leu Leu Leu Val Asp
                245                 250                 255

Asn Gly Cys His Gly His Pro Thr Val Glu Ala Lys Leu Arg Asn
                260                 265                 270

Gln Leu Glu Lys Tyr Ile Ser Glu Arg Thr Ile Gln Asp Ser Asn
                275                 280                 285

Tyr Gly Gly Lys Ile Pro Ile Val Cys Phe Ala Gln Gly Gly Gly
                290                 295                 300

Lys Glu Thr Leu Lys Ala Ile Asn Thr Ser Ile Lys Asn Lys Ile
                305                 310                 315

Pro Cys Val Val Val Glu Gly Ser Gly Gln Ile Ala Asp Val Ile
                320                 325                 330
```

```
Ala Ser Leu Val Glu Val Glu Asp Ala Leu Thr Ser Ser Ala Val
                335                 340                 345

Lys Glu Lys Leu Val Arg Phe Leu Pro Arg Thr Val Ser Arg Leu
                350                 355                 360

Pro Glu Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu Lys Glu Ile
                365                 370                 375

Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu Glu Ala
                380                 385                 390

Gly Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr Lys
                395                 400                 405

Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln
                410                 415                 420

Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp
                425                 430                 435

Glu Ile Phe Thr Asn Asp Arg Arg Trp Glu Ser Ala Asp Leu Gln
                440                 445                 450

Glu Val Met Phe Thr Ala Leu Ile Lys Asp Arg Pro Lys Phe Val
                455                 460                 465

Arg Leu Phe Leu Glu Asn Gly Leu Asn Leu Arg Lys Phe Leu Thr
                470                 475                 480

His Asp Val Leu Thr Glu Leu Phe Ser Asn His Phe Ser Thr Leu
                485                 490                 495

Val Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser Tyr Asn Asp Ala
                500                 505                 510

Leu Leu Thr Phe Val Trp Lys Leu Val Ala Asn Phe Arg Arg Gly
                515                 520                 525

Phe Arg Lys Glu Asp Arg Asn Gly Arg Asp Glu Met Asp Ile Glu
                530                 535                 540

Leu His Asp Val Ser Pro Ile Thr Arg His Pro Leu Gln Ala Leu
                545                 550                 555

Phe Ile Trp Ala Ile Leu Gln Asn Lys Lys Glu Leu Ser Lys Val
                560                 565                 570

Ile Trp Glu Gln Thr Arg Gly Cys Thr Leu Ala Ala Leu Gly Ala
                575                 580                 585

Ser Lys Leu Leu Lys Thr Leu Ala Lys Val Lys Asn Asp Ile Asn
                590                 595                 600

Ala Ala Gly Glu Ser Glu Glu Leu Ala Asn Glu Tyr Glu Thr Arg
                605                 610                 615

Ala Val Glu Leu Phe Thr Glu Cys Tyr Ser Ser Asp Glu Asp Leu
                620                 625                 630
```

```
Ala Glu Gln Leu Leu Val Tyr Ser Cys Glu Ala Trp Gly Gly Ser
                635                 640                 645

Asn Cys Leu Glu Leu Ala Val Glu Ala Thr Asp Gln His Phe Ile
                650                 655                 660

Ala Gln Pro Gly Val Gln Asn Phe Leu Ser Lys Gln Trp Tyr Gly
                665                 670                 675

Glu Ile Ser Arg Asp Thr Lys Asn Trp Lys Ile Ile Leu Cys Leu
                680                 685                 690

Phe Ile Ile Pro Leu Val Gly Cys Gly Phe Val Ser Phe Arg Lys
                695                 700                 705

Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp Tyr Tyr Val Ala
                710                 715                 720

Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn Val Val Phe
                725                 730                 735

Tyr Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu Met Asp
                740                 745                 750

Phe His Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser Leu
                755                 760                 765

Val Phe Val Leu Phe Cys Asp Glu Val Arg Gln Trp Tyr Val Asn
                770                 775                 780

Gly Val Asn Tyr Phe Thr Asp Leu Trp Asn Val Met Asp Thr Leu
                785                 790                 795

Gly Leu Phe Tyr Phe Ile Ala Gly Ile Val Phe Arg Leu His Ser
                800                 805                 810

Ser Asn Lys Ser Ser Leu Tyr Ser Gly Arg Val Ile Phe Cys Leu
                815                 820                 825

Asp Tyr Ile Ile Phe Thr Leu Arg Leu Ile His Ile Phe Thr Val
                830                 835                 840

Ser Arg Asn Leu Gly Pro Lys Ile Ile Met Leu Gln Arg Met Leu
                845                 850                 855

Ile Asp Val Phe Phe Phe Leu Phe Leu Phe Ala Val Trp Met Val
                860                 865                 870

Ala Phe Gly Val Ala Arg Gln Gly Ile Leu Arg Gln Asn Glu Gln
                875                 880                 885

Arg Trp Arg Trp Ile Phe Arg Ser Val Ile Tyr Glu Pro Tyr Leu
                890                 895                 900

Ala Met Phe Gly Gln Val Pro Ser Asp Val Asp Gly Thr Thr Tyr
                905                 910                 915

Asp Phe Ala His Cys Thr Phe Thr Gly Asn Glu Ser Lys Pro Leu
                920                 925                 930

Cys Val Glu Leu Asp Glu His Asn Leu Pro Arg Phe Pro Glu Trp
```

```
                935                 940                 945

 Ile Thr Ile Pro Leu Val Cys Ile Tyr Met Leu Ser Thr Asn Ile
                950                 955                 960

 Leu Leu Val Asn Leu Leu Val Ala Met Phe Gly Tyr Thr Val Gly
                965                 970                 975

 Thr Val Gln Glu Asn Asn Asp Gln Val Trp Lys Phe Gln Arg Tyr
                980                 985                 990

 Phe Leu Val Gln Glu Tyr Cys Ser Arg Leu Asn Ile Pro Phe Pro
                995                1000                1005

 Phe Ile Val Phe Ala Tyr Phe Tyr Met Val Val Lys Lys Cys Phe
               1010                1015                1020

 Lys Cys Cys Cys Lys Glu Lys Asn Met Glu Ser Ser Val Cys Cys
               1025                1030                1035

 Phe Lys Asn Glu Asp Asn Glu Thr Leu Ala Trp Glu Gly Val Met
               1040                1045                1050

 Lys Glu Asn Tyr Leu Val Lys Ile Asn Thr Lys Ala Asn Asp Thr
               1055                1060                1065

 Ser Glu Glu Met Arg His Arg Phe Arg Gln Leu Asp Thr Lys Leu
               1070                1075                1080

 Asn Asp Leu Lys Gly Leu Leu Lys Glu Ile Ala Asn Lys Ile Lys
               1085                1090                1095

<210> 97
<211> 347
<212> PRT
<213> Homo sapiens

<400> 97
 Arg Arg Gly Arg Gly Glu Val Leu Ala Val Glu Arg Gly Ser Gly
   1               5                  10                  15

 Ser Gly Gly Gly Gly Thr Arg Ser Gly Trp Pro Ala Pro Ala Ala
                20                  25                  30

 Gly Ala Asp Lys Lys Pro Ser Arg Cys Gly Ser Gly Arg Glu Gly
                35                  40                  45

 Glu Gly Val Ser Glu Gly His Lys Ser Met Thr Gly Leu Tyr Glu
                50                  55                  60

 Leu Val Trp Arg Val Leu His Ala Leu Leu Cys Leu His Arg Thr
                65                  70                  75

 Leu Thr Ser Trp Leu Arg Val Arg Phe Gly Thr Trp Asn Trp Ile
                80                  85                  90

 Trp Arg Arg Cys Cys Arg Ala Ala Ser Ala Ala Val Leu Ala Pro
                95                 100                 105

 Leu Gly Phe Thr Leu Arg Lys Pro Pro Ala Val Gly Arg Asn Arg
               110                 115                 120
```

```
 Arg His His Arg His Pro Arg Gly Gly Ser Cys Leu Ala Ala Ala
                 125                 130                 135

 His His Arg Met Arg Trp Arg Ala Asp Gly Arg Ser Leu Glu Lys
                 140                 145                 150

 Leu Pro Val His Met Gly Leu Val Ile Thr Glu Val Glu Gln Glu
                 155                 160                 165

 Pro Ser Phe Ser Asp Ile Ala Ser Leu Val Val Trp Cys Met Ala
                 170                 175                 180

 Val Gly Ile Ser Tyr Ile Ser Val Tyr Asp His Gln Gly Ile Phe
                 185                 190                 195

 Lys Arg Asn Asn Ser Arg Leu Met Asp Glu Ile Leu Lys Gln Gln
                 200                 205                 210

 Gln Glu Leu Leu Gly Leu Asp Cys Ser Lys Tyr Ser Pro Glu Phe
                 215                 220                 225

 Ala Asn Ser Asn Asp Lys Asp Asp Gln Val Leu Asn Cys His Leu
                 230                 235                 240

 Ala Val Lys Val Leu Ser Pro Glu Asp Gly Lys Ala Asp Ile Val
                 245                 250                 255

 Arg Ala Ala Gln Asp Phe Cys Gln Leu Val Ala Gln Lys Gln Arg
                 260                 265                 270

 Arg Pro Thr Asp Leu Asp Val Asp Thr Leu Ala Ser Leu Leu Ser
                 275                 280                 285

 Ser Asn Gly Cys Pro Asp Pro Asp Leu Val Leu Lys Phe Gly Pro
                 290                 295                 300

 Val Asp Ser Thr Leu Gly Phe Leu Pro Trp His Ile Arg Leu Thr
                 305                 310                 315

 Glu Ile Val Ser Leu Pro Ser His Leu Asn Ile Ser Tyr Glu Asp
                 320                 325                 330

 Phe Phe Ser Ala Leu Arg Gln Tyr Ala Ala Cys Glu Gln Arg Leu
                 335                 340                 345

 Gly Lys


<210> 98
<211> 293
<212> PRT
<213> Homo sapiens

<400> 98
 Met Thr Gly Leu Tyr Glu Leu Val Trp Arg Val Leu His Ala Leu
   1               5                  10                  15

 Leu Cys Leu His Arg Thr Leu Thr Ser Trp Leu Arg Val Arg Phe
                  20                  25                  30
```

```
 Gly Thr Trp Asn Trp Ile Trp Arg Arg Cys Cys Arg Ala Ala Ser
                  35                  40                  45

 Ala Ala Val Leu Ala Pro Leu Gly Phe Thr Leu Arg Lys Pro Pro
                  50                  55                  60

 Ala Val Gly Arg Asn Arg Arg His His Arg His Pro Arg Gly Gly
                  65                  70                  75

 Ser Cys Leu Ala Ala Ala His His Arg Met Arg Trp Arg Ala Asp
                  80                  85                  90

 Gly Arg Ser Leu Glu Lys Leu Pro Val His Met Gly Leu Val Ile
                  95                 100                 105

 Thr Glu Val Glu Gln Glu Pro Ser Phe Ser Asp Ile Ala Ser Leu
                 110                 115                 120

 Val Val Trp Cys Met Ala Val Gly Ile Ser Tyr Ile Ser Val Tyr
                 125                 130                 135

 Asp His Gln Gly Ile Phe Lys Arg Asn Asn Ser Arg Leu Met Asp
                 140                 145                 150

 Glu Ile Leu Lys Gln Gln Gln Glu Leu Leu Gly Leu Asp Cys Ser
                 155                 160                 165

 Lys Tyr Ser Pro Glu Phe Ala Asn Ser Asn Asp Lys Asp Asp Gln
                 170                 175                 180

 Val Leu Asn Cys His Leu Ala Val Lys Val Leu Ser Pro Glu Asp
                 185                 190                 195

 Gly Lys Ala Asp Ile Val Arg Ala Ala Gln Asp Phe Cys Gln Leu
                 200                 205                 210

 Val Ala Gln Lys Gln Lys Arg Pro Thr Asp Leu Asp Val Asp Thr
                 215                 220                 225

 Leu Ala Ser Leu Leu Ser Ser Asn Gly Cys Pro Asp Pro Asp Leu
                 230                 235                 240

 Val Leu Lys Phe Gly Pro Val Asp Ser Thr Leu Gly Phe Leu Pro
                 245                 250                 255

 Trp His Ile Arg Leu Thr Glu Ile Val Ser Leu Pro Ser His Leu
                 260                 265                 270

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

 Ala Cys Glu Gln Arg Leu Gly Lys
                 290

<210> 99
<211> 553
<212> PRT
<213> Homo sapiens

<400> 99
 Met Val Gln Arg Leu Trp Val Ser Arg Leu Leu Arg His Arg Lys
```

```
  1               5                  10                  15

Ala Gln Leu Leu Leu Val Asn Leu Leu Thr Phe Gly Leu Glu Val
                20                  25                  30

Cys Leu Ala Ala Gly Ile Thr Tyr Val Pro Pro Leu Leu Leu Glu
                35                  40                  45

Val Gly Val Glu Glu Lys Phe Met Thr Met Val Leu Gly Ile Gly
                50                  55                  60

Pro Val Leu Gly Leu Val Cys Val Pro Leu Leu Gly Ser Ala Ser
                65                  70                  75

Asp His Trp Arg Gly Arg Tyr Gly Arg Arg Arg Pro Phe Ile Trp
                80                  85                  90

Ala Leu Ser Leu Gly Ile Leu Leu Ser Leu Phe Leu Ile Pro Arg
                95                 100                 105

Ala Gly Trp Leu Ala Gly Leu Leu Cys Pro Asp Pro Arg Pro Leu
               110                 115                 120

Glu Leu Ala Leu Leu Ile Leu Gly Val Gly Leu Leu Asp Phe Cys
               125                 130                 135

Gly Gln Val Cys Phe Thr Pro Leu Glu Ala Leu Leu Ser Asp Leu
               140                 145                 150

Phe Arg Asp Pro Asp His Cys Arg Gln Ala Tyr Ser Val Tyr Ala
               155                 160                 165

Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr Leu Leu Pro Ala
               170                 175                 180

Ile Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu Gly Thr Gln
               185                 190                 195

Glu Glu Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu Thr Cys
               200                 205                 210

Val Ala Ala Thr Leu Leu Val Ala Glu Glu Ala Ala Leu Gly Pro
               215                 220                 225

Thr Glu Pro Ala Glu Gly Leu Ser Ala Pro Ser Leu Ser Pro His
               230                 235                 240

Cys Cys Pro Cys Arg Ala Arg Leu Ala Phe Arg Asn Leu Gly Ala
               245                 250                 255

Leu Leu Pro Arg Leu His Gln Leu Cys Cys Arg Met Pro Arg Thr
               260                 265                 270

Leu Arg Arg Leu Phe Val Ala Glu Leu Cys Ser Trp Met Ala Leu
               275                 280                 285

Met Thr Phe Thr Leu Phe Tyr Thr Asp Phe Val Gly Glu Gly Leu
               290                 295                 300

Tyr Gln Gly Val Pro Arg Ala Glu Pro Gly Thr Glu Ala Arg Arg
               305                 310                 315
```

```
 His Tyr Asp Glu Gly Val Arg Met Gly Ser Leu Gly Leu Phe Leu
                 320                 325                 330

 Gln Cys Ala Ile Ser Leu Val Phe Ser Leu Val Met Asp Arg Leu
                 335                 340                 345

 Val Gln Arg Phe Gly Thr Arg Ala Val Tyr Leu Ala Ser Val Ala
                 350                 355                 360

 Ala Phe Pro Val Ala Ala Gly Ala Thr Cys Leu Ser His Ser Val
                 365                 370                 375

 Ala Val Val Thr Ala Ser Ala Ala Leu Thr Gly Phe Thr Phe Ser
                 380                 385                 390

 Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala Ser Leu Tyr His Arg
                 395                 400                 405

 Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly Asp Thr Gly Gly
                 410                 415                 420

 Ala Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu Pro Gly Pro
                 425                 430                 435

 Lys Pro Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly
                 440                 445                 450

 Ser Gly Leu Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala
                 455                 460                 465

 Cys Asp Val Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala
                 470                 475                 480

 Arg Val Val Pro Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu
                 485                 490                 495

 Asp Ser Ala Phe Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met
                 500                 505                 510

 Gly Ser Ile Val Gln Leu Ser Gln Ser Val Thr Ala Tyr Met Val
                 515                 520                 525

 Ser Ala Ala Gly Leu Gly Leu Val Ala Ile Tyr Phe Ala Thr Gln
                 530                 535                 540

 Val Val Phe Asp Lys Ser Asp Leu Ala Lys Tyr Ser Ala
                 545                 550

<210> 100
<211> 987
<212> PRT
<213> Homo sapiens

<400> 100
 Met Ala Leu Arg Arg Leu Gly Ala Ala Leu Leu Leu Leu Pro Leu
   1               5                  10                  15

 Leu Ala Ala Val Glu Glu Thr Leu Met Asp Ser Thr Thr Ala Thr
                  20                  25                  30
```

```
Ala Glu Leu Gly Trp Met Val His Pro Pro Ser Gly Trp Glu Glu
                 35                  40                  45

Val Ser Gly Tyr Asp Glu Asn Met Asn Thr Ile Arg Thr Tyr Gln
                 50                  55                  60

Val Cys Asn Val Phe Glu Ser Ser Gln Asn Asn Trp Leu Arg Thr
                 65                  70                  75

Lys Phe Ile Arg Arg Arg Gly Ala His Arg Ile His Val Glu Met
                 80                  85                  90

Lys Phe Ser Val Arg Asp Cys Ser Ser Ile Pro Ser Val Pro Gly
                 95                 100                 105

Ser Cys Lys Glu Thr Phe Asn Leu Tyr Tyr Tyr Glu Ala Asp Phe
                110                 115                 120

Asp Ser Ala Thr Lys Thr Phe Pro Asn Trp Met Glu Asn Pro Trp
                125                 130                 135

Val Lys Val Asp Thr Ile Ala Ala Asp Glu Ser Phe Ser Gln Val
                140                 145                 150

Asp Leu Gly Gly Arg Val Met Lys Ile Asn Thr Glu Val Arg Ser
                155                 160                 165

Phe Gly Pro Val Ser Arg Ser Gly Phe Tyr Leu Ala Phe Gln Asp
                170                 175                 180

Tyr Gly Gly Cys Met Ser Leu Ile Ala Val Arg Val Phe Tyr Arg
                185                 190                 195

Lys Cys Pro Arg Ile Ile Gln Asn Gly Ala Ile Phe Gln Glu Thr
                200                 205                 210

Leu Ser Gly Ala Glu Ser Thr Ser Leu Val Ala Ala Arg Gly Ser
                215                 220                 225

Cys Ile Ala Asn Ala Glu Glu Val Asp Val Pro Ile Lys Leu Tyr
                230                 235                 240

Cys Asn Gly Asp Gly Glu Trp Leu Val Pro Ile Gly Arg Cys Met
                245                 250                 255

Cys Lys Ala Gly Phe Glu Ala Val Glu Asn Gly Thr Val Cys Arg
                260                 265                 270

Gly Cys Pro Ser Gly Thr Phe Lys Ala Asn Gln Gly Asp Glu Ala
                275                 280                 285

Cys Thr His Cys Pro Ile Asn Ser Arg Thr Thr Ser Glu Gly Ala
                290                 295                 300

Thr Asn Cys Val Cys Arg Asn Gly Tyr Tyr Arg Ala Asp Leu Asp
                305                 310                 315

Pro Leu Asp Met Pro Cys Thr Thr Ile Pro Ser Ala Pro Gln Ala
                320                 325                 330

Val Ile Ser Ser Val Asn Glu Thr Ser Leu Met Leu Glu Trp Thr
```

```
                335                 340                 345

Pro Pro Arg Asp Ser Gly Gly Arg Glu Asp Leu Val Tyr Asn Ile
                350                 355                 360

Ile Cys Lys Ser Cys Gly Ser Gly Arg Gly Ala Cys Thr Arg Cys
                365                 370                 375

Gly Asp Asn Val Gln Tyr Ala Pro Arg Gln Leu Gly Leu Thr Glu
                380                 385                 390

Pro Arg Ile Tyr Ile Ser Asp Leu Leu Ala His Thr Gln Tyr Thr
                395                 400                 405

Phe Glu Ile Gln Ala Val Asn Gly Val Thr Asp Gln Ser Pro Phe
                410                 415                 420

Ser Pro Gln Phe Ala Ser Val Asn Ile Thr Thr Asn Gln Ala Ala
                425                 430                 435

Pro Ser Ala Val Ser Ile Met His Gln Val Ser Arg Thr Val Asp
                440                 445                 450

Ser Ile Thr Leu Ser Trp Ser Gln Pro Asp Gln Pro Asn Gly Val
                455                 460                 465

Ile Leu Asp Tyr Glu Leu Gln Tyr Tyr Glu Lys Glu Leu Ser Glu
                470                 475                 480

Tyr Asn Ala Thr Ala Ile Lys Ser Pro Thr Asn Thr Val Thr Val
                485                 490                 495

Gln Gly Leu Lys Ala Gly Ala Ile Tyr Val Phe Gln Val Arg Ala
                500                 505                 510

Arg Thr Val Ala Gly Tyr Gly Arg Tyr Ser Gly Lys Met Tyr Phe
                515                 520                 525

Gln Thr Met Thr Glu Ala Glu Tyr Gln Thr Ser Ile Gln Glu Lys
                530                 535                 540

Leu Pro Leu Ile Ile Gly Ser Ser Ala Ala Gly Leu Val Phe Leu
                545                 550                 555

Ile Ala Val Val Val Ile Ala Ile Val Cys Asn Arg Arg Arg Gly
                560                 565                 570

Phe Glu Arg Ala Asp Ser Glu Tyr Thr Asp Lys Leu Gln His Tyr
                575                 580                 585

Thr Ser Gly His Met Thr Pro Gly Met Lys Ile Tyr Ile Asp Pro
                590                 595                 600

Phe Thr Tyr Glu Asp Pro Asn Glu Ala Val Arg Glu Phe Ala Lys
                605                 610                 615

Glu Ile Asp Ile Ser Cys Val Lys Ile Glu Gln Val Ile Gly Ala
                620                 625                 630

Gly Glu Phe Gly Glu Val Cys Ser Gly His Leu Lys Leu Pro Gly
                635                 640                 645
```

```
Lys Arg Glu Ile Phe Val Ala Ile Lys Thr Leu Lys Ser Gly Tyr
                650                 655                 660

Thr Glu Lys Gln Arg Arg Asp Phe Leu Ser Glu Ala Ser Ile Met
                665                 670                 675

Gly Gln Phe Asp His Pro Asn Val Ile His Leu Glu Gly Val Val
                680                 685                 690

Thr Lys Ser Thr Pro Val Met Ile Ile Thr Glu Phe Met Glu Asn
                695                 700                 705

Gly Ser Leu Asp Ser Phe Leu Arg Gln Asn Asp Gly Gln Phe Thr
                710                 715                 720

Val Ile Gln Leu Val Gly Met Leu Arg Gly Ile Ala Ala Gly Met
                725                 730                 735

Lys Tyr Leu Ala Asp Met Asn Tyr Val His Arg Asp Leu Ala Ala
                740                 745                 750

Arg Asn Ile Leu Val Asn Ser Asn Leu Val Cys Lys Val Ser Asp
                755                 760                 765

Phe Gly Leu Ser Arg Phe Leu Glu Asp Asp Thr Ser Asp Pro Thr
                770                 775                 780

Tyr Thr Ser Ala Leu Gly Gly Lys Ile Pro Ile Arg Trp Thr Ala
                785                 790                 795

Pro Glu Ala Ile Gln Tyr Arg Lys Phe Thr Ser Ala Ser Asp Val
                800                 805                 810

Trp Ser Tyr Gly Ile Val Met Trp Glu Val Met Ser Tyr Gly Glu
                815                 820                 825

Arg Pro Tyr Trp Asp Met Thr Asn Gln Asp Val Ile Asn Ala Ile
                830                 835                 840

Glu Gln Asp Tyr Arg Leu Pro Pro Pro Met Asp Cys Pro Ser Ala
                845                 850                 855

Leu His Gln Leu Met Leu Asp Cys Trp Gln Lys Asp Arg Asn His
                860                 865                 870

Arg Pro Lys Phe Gly Gln Ile Val Asn Thr Leu Asp Lys Met Ile
                875                 880                 885

Arg Asn Pro Asn Ser Leu Lys Ala Met Ala Pro Leu Ser Ser Gly
                890                 895                 900

Ile Asn Leu Pro Leu Leu Asp Arg Thr Ile Pro Asp Tyr Thr Ser
                905                 910                 915

Phe Asn Thr Val Asp Glu Trp Leu Glu Ala Ile Lys Met Gly Gln
                920                 925                 930

Tyr Lys Glu Ser Phe Ala Asn Ala Gly Phe Thr Ser Phe Asp Val
                935                 940                 945
```

```
 Val Ser Gln Met Met Met Glu Asp Ile Leu Arg Val Gly Leu Thr
                 950                 955                 960

 Leu Ala Gly His Gln Lys Lys Ile Leu Asn Ser Ile Gln Val Met
                 965                 970                 975

 Arg Ala Gln Met Asn Gln Ile Gln Ser Val Glu Val
                 980                 985

<210> 101
<211> 986
<212> PRT
<213> Homo sapiens

<400> 101
 Met Ala Leu Arg Arg Leu Gly Ala Ala Leu Leu Leu Leu Pro Leu
   1               5                  10                  15

 Leu Ala Ala Val Glu Glu Thr Leu Met Asp Ser Thr Thr Ala Thr
                  20                  25                  30

 Ala Glu Leu Gly Trp Met Val His Pro Pro Ser Gly Trp Glu Glu
                  35                  40                  45

 Val Ser Gly Tyr Asp Glu Asn Met Asn Thr Ile Arg Thr Tyr Gln
                  50                  55                  60

 Val Cys Asn Val Phe Glu Ser Ser Gln Asn Asn Trp Leu Arg Thr
                  65                  70                  75

 Lys Phe Ile Arg Arg Arg Gly Ala His Arg Ile His Val Glu Met
                  80                  85                  90

 Lys Phe Ser Val Arg Asp Cys Ser Ser Ile Pro Ser Val Pro Gly
                  95                 100                 105

 Ser Cys Lys Glu Thr Phe Asn Leu Tyr Tyr Tyr Glu Ala Asp Phe
                 110                 115                 120

 Asp Ser Ala Thr Lys Thr Phe Pro Asn Trp Met Glu Asn Pro Trp
                 125                 130                 135

 Val Lys Val Asp Thr Ile Ala Ala Asp Glu Ser Phe Ser Gln Val
                 140                 145                 150

 Asp Leu Gly Gly Arg Val Met Lys Ile Asn Thr Glu Val Arg Ser
                 155                 160                 165

 Phe Gly Pro Val Ser Arg Ser Gly Phe Tyr Leu Ala Phe Gln Asp
                 170                 175                 180

 Tyr Gly Gly Cys Met Ser Leu Ile Ala Val Arg Val Phe Tyr Arg
                 185                 190                 195

 Lys Cys Pro Arg Ile Ile Gln Asn Gly Ala Ile Phe Gln Glu Thr
                 200                 205                 210

 Leu Ser Gly Ala Glu Ser Thr Ser Leu Val Ala Ala Arg Gly Ser
                 215                 220                 225

 Cys Ile Ala Asn Ala Glu Glu Val Asp Val Pro Ile Lys Leu Tyr
```

```
                230                 235                 240

Cys Asn Gly Asp Gly Glu Trp Leu Val Pro Ile Gly Arg Cys Met
                245                 250                 255

Cys Lys Ala Gly Phe Glu Ala Val Glu Asn Gly Thr Val Cys Arg
                260                 265                 270

Gly Cys Pro Ser Gly Thr Phe Lys Ala Asn Gln Gly Asp Glu Ala
                275                 280                 285

Cys Thr His Cys Pro Ile Asn Ser Arg Thr Thr Ser Glu Gly Ala
                290                 295                 300

Thr Asn Cys Val Cys Arg Asn Gly Tyr Tyr Arg Ala Asp Leu Asp
                305                 310                 315

Pro Leu Asp Met Pro Cys Thr Thr Ile Pro Ser Ala Pro Gln Ala
                320                 325                 330

Val Ile Ser Ser Val Asn Glu Thr Ser Leu Met Leu Glu Trp Thr
                335                 340                 345

Pro Pro Arg Asp Ser Gly Gly Arg Glu Asp Leu Val Tyr Asn Ile
                350                 355                 360

Ile Cys Lys Ser Cys Gly Ser Gly Arg Gly Ala Cys Thr Arg Cys
                365                 370                 375

Gly Asp Asn Val Gln Tyr Ala Pro Arg Gln Leu Gly Leu Thr Glu
                380                 385                 390

Pro Arg Ile Tyr Ile Ser Asp Leu Leu Ala His Thr Gln Tyr Thr
                395                 400                 405

Phe Glu Ile Gln Ala Val Asn Gly Val Thr Asp Gln Ser Pro Phe
                410                 415                 420

Ser Pro Gln Phe Ala Ser Val Asn Ile Thr Thr Asn Gln Ala Ala
                425                 430                 435

Pro Ser Ala Val Ser Ile Met His Gln Val Ser Arg Thr Val Asp
                440                 445                 450

Ser Ile Thr Leu Ser Trp Ser Gln Pro Asp Gln Pro Asn Gly Val
                455                 460                 465

Ile Leu Asp Tyr Glu Leu Gln Tyr Tyr Glu Lys Glu Leu Ser Glu
                470                 475                 480

Tyr Asn Ala Thr Ala Ile Lys Ser Pro Thr Asn Thr Val Thr Val
                485                 490                 495

Gln Gly Leu Lys Ala Gly Ala Ile Tyr Val Phe Gln Val Arg Ala
                500                 505                 510

Arg Thr Val Ala Gly Tyr Gly Arg Tyr Ser Gly Lys Met Tyr Phe
                515                 520                 525

Gln Thr Met Thr Glu Ala Glu Tyr Gln Thr Ser Ile Gln Glu Lys
                530                 535                 540
```

```
Leu Pro Leu Ile Ile Gly Ser Ser Ala Ala Gly Leu Val Phe Leu
                545                 550                 555

Ile Ala Val Val Val Ile Ala Ile Val Cys Asn Arg Arg Gly Phe
                560                 565                 570

Glu Arg Ala Asp Ser Glu Tyr Thr Asp Lys Leu Gln His Tyr Thr
                575                 580                 585

Ser Gly His Met Thr Pro Gly Met Lys Ile Tyr Ile Asp Pro Phe
                590                 595                 600

Thr Tyr Glu Asp Pro Asn Glu Ala Val Arg Glu Phe Ala Lys Glu
                605                 610                 615

Ile Asp Ile Ser Cys Val Lys Ile Glu Gln Val Ile Gly Ala Gly
                620                 625                 630

Glu Phe Gly Glu Val Cys Ser Gly His Leu Lys Leu Pro Gly Lys
                635                 640                 645

Arg Glu Ile Phe Val Ala Ile Lys Thr Leu Lys Ser Gly Tyr Thr
                650                 655                 660

Glu Lys Gln Arg Arg Asp Phe Leu Ser Glu Ala Ser Ile Met Gly
                665                 670                 675

Gln Phe Asp His Pro Asn Val Ile His Leu Glu Gly Val Val Thr
                680                 685                 690

Lys Ser Thr Pro Val Met Ile Ile Thr Glu Phe Met Glu Asn Gly
                695                 700                 705

Ser Leu Asp Ser Phe Leu Arg Gln Asn Asp Gly Gln Phe Thr Val
                710                 715                 720

Ile Gln Leu Val Gly Met Leu Arg Gly Ile Ala Ala Gly Met Lys
                725                 730                 735

Tyr Leu Ala Asp Met Asn Tyr Val His Arg Asp Leu Ala Ala Arg
                740                 745                 750

Asn Ile Leu Val Asn Ser Asn Leu Val Cys Lys Val Ser Asp Phe
                755                 760                 765

Gly Leu Ser Arg Phe Leu Glu Asp Asp Thr Ser Asp Pro Thr Tyr
                770                 775                 780

Thr Ser Ala Leu Gly Gly Lys Ile Pro Ile Arg Trp Thr Ala Pro
                785                 790                 795

Glu Ala Ile Gln Tyr Arg Lys Phe Thr Ser Ala Ser Asp Val Trp
                800                 805                 810

Ser Tyr Gly Ile Val Met Trp Glu Val Met Ser Tyr Gly Glu Arg
                815                 820                 825

Pro Tyr Trp Asp Met Thr Asn Gln Asp Val Ile Asn Ala Ile Glu
                830                 835                 840
```

```
 Gln Asp Tyr Arg Leu Pro Pro Pro Met Asp Cys Pro Ser Ala Leu
                 845                 850                 855

 His Gln Leu Met Leu Asp Cys Trp Gln Lys Asp Arg Asn His Arg
                 860                 865                 870

 Pro Lys Phe Gly Gln Ile Val Asn Thr Leu Asp Lys Met Ile Arg
                 875                 880                 885

 Asn Pro Asn Ser Leu Lys Ala Met Ala Pro Leu Ser Ser Gly Ile
                 890                 895                 900

 Asn Leu Pro Leu Leu Asp Arg Thr Ile Pro Asp Tyr Thr Ser Phe
                 905                 910                 915

 Asn Thr Val Asp Glu Trp Leu Glu Ala Ile Lys Met Gly Gln Tyr
                 920                 925                 930

 Lys Glu Ser Phe Ala Asn Ala Gly Phe Thr Ser Phe Asp Val Val
                 935                 940                 945

 Ser Gln Met Met Met Glu Asp Ile Leu Arg Val Gly Val Thr Leu
                 950                 955                 960

 Ala Gly His Gln Lys Lys Ile Leu Asn Ser Ile Gln Val Met Arg
                 965                 970                 975

 Ala Gln Met Asn Gln Ile Gln Ser Val Glu Val
                 980                 985

<210> 102
<211> 153
<212> PRT
<213> Homo sapiens

<400> 102
 Met Asn Phe Ser Thr Ser Ser Ser Ser Phe Ala Tyr Asp Arg Glu
   1               5                  10                  15

 Phe Leu Arg Thr Leu Pro Gly Phe Leu Ile Val Ala Glu Ile Val
                  20                  25                  30

 Leu Gly Leu Leu Val Trp Thr Leu Ile Ala Gly Thr Glu Tyr Phe
                  35                  40                  45

 Arg Val Pro Ala Phe Gly Trp Val Met Phe Val Ala Val Phe Tyr
                  50                  55                  60

 Trp Val Leu Thr Val Phe Phe Leu Ile Ile Tyr Ile Thr Met Thr
                  65                  70                  75

 Tyr Thr Arg Ile Pro Gln Val Pro Trp Thr Thr Val Gly Leu Cys
                  80                  85                  90

 Phe Asn Gly Ser Ala Phe Val Leu Tyr Leu Ser Ala Ala Val Val
                  95                 100                 105

 Asp Ala Ser Ser Val Ser Pro Glu Arg Asp Ser His Asn Phe Asn
                 110                 115                 120

 Ser Trp Ala Ala Ser Ser Phe Phe Ala Phe Leu Val Asn Ile Cys
```

```
                125                 130                 135

 Tyr Ala Gly Asn Thr Tyr Phe Ser Phe Ile Ala Trp Arg Ser Arg
                140                 145                 150

 Thr Ile Gln


<210> 103
<211> 153
<212> PRT
<213> Homo sapiens

<400> 103
 Met Asn Phe Ser Thr Ser Ser Ser Ser Phe Ala Tyr Asp Arg Glu
   1               5                  10                  15

 Phe Leu Arg Thr Leu Pro Gly Phe Leu Ile Val Ala Glu Ile Val
                20                  25                  30

 Leu Gly Leu Leu Val Trp Thr Leu Ile Ala Gly Thr Glu Tyr Phe
                35                  40                  45

 Arg Val Pro Ala Phe Gly Trp Val Met Phe Val Ala Val Phe Tyr
                50                  55                  60

 Trp Val Leu Thr Val Phe Phe Leu Ile Ile Tyr Ile Thr Met Thr
                65                  70                  75

 Tyr Thr Arg Ile Pro Gln Val Pro Trp Thr Thr Val Gly Leu Cys
                80                  85                  90

 Phe Asn Gly Ser Ala Phe Val Leu Tyr Leu Ser Ala Ala Val Val
                95                 100                 105

 Asp Ala Ser Ser Val Ser Pro Glu Arg Asp Ser His Asn Phe Asn
               110                 115                 120

 Ser Trp Ala Ala Ser Ser Phe Phe Ala Phe Leu Val Thr Ile Cys
               125                 130                 135

 Tyr Ala Gly Asn Thr Tyr Phe Ser Phe Ile Ala Trp Arg Ser Arg
               140                 145                 150

 Thr Ile Gln


<210> 104
<211> 555
<212> PRT
<213> Homo sapiens

<400> 104
 Met Pro Thr Thr Val Asp Asp Val Leu Glu His Gly Gly Glu Phe
   1               5                  10                  15

 His Phe Phe Gln Lys Gln Met Phe Phe Leu Leu Ala Leu Leu Ser
                20                  25                  30

 Ala Thr Phe Ala Pro Ile Tyr Val Gly Ile Val Phe Leu Gly Phe
                35                  40                  45
```

```
Thr Pro Asp His Arg Cys Arg Ser Pro Gly Val Ala Glu Leu Ser
                 50                  55                  60

Leu Arg Cys Gly Trp Ser Pro Ala Glu Glu Leu Asn Tyr Thr Val
                 65                  70                  75

Pro Gly Pro Gly Pro Ala Gly Glu Ala Ser Pro Arg Gln Cys Arg
                 80                  85                  90

Arg Tyr Glu Val Asp Trp Asn Gln Ser Thr Phe Asp Cys Val Asp
                 95                 100                 105

Pro Leu Ala Ser Leu Asp Thr Asn Arg Ser Arg Leu Pro Leu Gly
                110                 115                 120

Pro Cys Arg Asp Gly Trp Val Tyr Glu Thr Pro Gly Ser Ser Ile
                125                 130                 135

Val Thr Glu Phe Asn Leu Val Cys Ala Asn Ser Trp Met Leu Asp
                140                 145                 150

Leu Phe Gln Ser Ser Val Asn Val Gly Phe Phe Ile Gly Ser Met
                155                 160                 165

Ser Ile Gly Tyr Ile Ala Asp Arg Phe Gly Arg Lys Leu Cys Leu
                170                 175                 180

Leu Thr Thr Val Leu Ile Asn Ala Ala Ala Gly Val Leu Met Ala
                185                 190                 195

Ile Ser Pro Thr Tyr Thr Trp Met Leu Ile Phe Arg Leu Ile Gln
                200                 205                 210

Gly Leu Val Ser Lys Ala Gly Trp Leu Ile Gly Tyr Ile Leu Ile
                215                 220                 225

Thr Glu Phe Val Gly Arg Arg Tyr Arg Arg Thr Val Gly Ile Phe
                230                 235                 240

Tyr Gln Val Ala Tyr Thr Val Gly Leu Leu Val Leu Ala Gly Val
                245                 250                 255

Ala Tyr Ala Leu Pro His Trp Arg Trp Leu Gln Phe Thr Val Ala
                260                 265                 270

Leu Pro Asn Phe Phe Phe Leu Leu Tyr Tyr Trp Cys Ile Pro Glu
                275                 280                 285

Ser Pro Arg Trp Leu Ile Ser Gln Asn Lys Asn Ala Glu Ala Met
                290                 295                 300

Arg Ile Ile Lys His Ile Ala Lys Lys Asn Gly Lys Ser Leu Pro
                305                 310                 315

Ala Ser Leu Gln Arg Leu Arg Leu Glu Glu Glu Thr Gly Lys Lys
                320                 325                 330

Leu Asn Pro Ser Phe Leu Asp Leu Val Arg Thr Pro Gln Ile Arg
                335                 340                 345
```

```
 Lys His Thr Met Ile Leu Met Tyr Asn Trp Phe Thr Ser Ser Val
                 350                 355                 360

 Leu Tyr Gln Gly Leu Ile Met His Met Gly Leu Ala Gly Asp Asn
                 365                 370                 375

 Ile Tyr Leu Asp Phe Phe Tyr Ser Ala Leu Val Glu Phe Pro Ala
                 380                 385                 390

 Ala Phe Met Ile Ile Leu Thr Ile Asp Arg Ile Gly Arg Arg Tyr
                 395                 400                 405

 Pro Trp Ala Ala Ser Asn Met Val Ala Gly Ala Ala Cys Leu Ala
                 410                 415                 420

 Ser Val Phe Ile Pro Gly Asp Leu Gln Trp Leu Lys Ile Ile Ile
                 425                 430                 435

 Ser Cys Leu Gly Arg Met Gly Ile Thr Met Ala Tyr Glu Ile Val
                 440                 445                 450

 Cys Leu Val Asn Ala Glu Leu Tyr Pro Thr Phe Ile Arg Asn Leu
                 455                 460                 465

 Gly Val His Ile Cys Ser Ser Met Cys Asp Ile Gly Gly Ile Ile
                 470                 475                 480

 Thr Pro Phe Leu Val Tyr Arg Leu Thr Asn Ile Trp Leu Glu Leu
                 485                 490                 495

 Pro Leu Met Val Phe Gly Val Leu Gly Leu Val Ala Gly Gly Leu
                 500                 505                 510

 Val Leu Leu Leu Pro Glu Thr Lys Gly Lys Ala Leu Pro Glu Thr
                 515                 520                 525

 Ile Glu Glu Ala Glu Asn Met Gln Arg Pro Arg Lys Asn Lys Glu
                 530                 535                 540

 Lys Met Ile Tyr Leu Gln Val Gln Lys Leu Asp Ile Pro Leu Asn
                 545                 550                 555

<210> 105
<211> 555
<212> PRT
<213> Homo sapiens

<400> 105
 Met Pro Thr Thr Val Asp Asp Val Leu Glu His Gly Gly Glu Phe
   1               5                  10                  15

 His Phe Phe Gln Lys Gln Met Phe Phe Leu Leu Ala Leu Leu Ser
                  20                  25                  30

 Ala Thr Phe Ala Pro Ile Tyr Val Gly Ile Val Phe Leu Gly Phe
                  35                  40                  45

 Thr Pro Asp His Arg Cys Arg Ser Pro Gly Val Ala Glu Leu Ser
                  50                  55                  60

 Leu Arg Cys Gly Trp Ser Pro Ala Glu Glu Leu Asn Tyr Thr Val
```

```
                 65                  70                  75

Pro Gly Pro Gly Pro Ala Gly Glu Ala Ser Pro Arg Gln Cys Arg
                 80                  85                  90

Arg Tyr Glu Val Asp Trp Asn Gln Ser Thr Phe Asp Cys Val Asp
                 95                 100                 105

Pro Leu Ala Ser Leu Asp Thr Asn Arg Ser Arg Leu Pro Leu Gly
                110                 115                 120

Pro Cys Arg Asp Gly Trp Val Tyr Glu Thr Pro Gly Ser Ser Ile
                125                 130                 135

Val Thr Glu Phe Asn Leu Val Cys Ala Asn Ser Trp Met Leu Asp
                140                 145                 150

Leu Phe Gln Ser Ser Val Asn Val Gly Phe Phe Ile Gly Ser Met
                155                 160                 165

Ser Ile Gly Tyr Ile Ala Asp Arg Phe Gly Arg Lys Leu Cys Leu
                170                 175                 180

Leu Thr Thr Val Leu Ile Asn Ala Ala Ala Gly Val Leu Met Ala
                185                 190                 195

Ile Ser Pro Thr Tyr Thr Trp Met Leu Ile Phe Arg Leu Ile Gln
                200                 205                 210

Gly Leu Val Ser Lys Ala Gly Trp Leu Ile Gly Tyr Ile Leu Ile
                215                 220                 225

Thr Glu Phe Val Gly Arg Arg Tyr Arg Arg Thr Val Gly Ile Phe
                230                 235                 240

Tyr Gln Val Ala Tyr Thr Val Gly Leu Leu Val Leu Ala Gly Val
                245                 250                 255

Ala Tyr Ala Leu Pro His Trp Arg Trp Leu Gln Phe Thr Val Ser
                260                 265                 270

Leu Pro Asn Phe Phe Phe Leu Leu Tyr Tyr Trp Cys Ile Pro Glu
                275                 280                 285

Ser Pro Arg Trp Leu Ile Ser Gln Asn Lys Asn Ala Glu Ala Met
                290                 295                 300

Arg Ile Ile Lys His Ile Ala Lys Lys Asn Gly Lys Ser Leu Pro
                305                 310                 315

Ala Ser Leu Gln Arg Leu Arg Leu Glu Glu Glu Thr Gly Lys Lys
                320                 325                 330

Leu Asn Pro Ser Phe Leu Asp Leu Val Arg Thr Pro Gln Ile Arg
                335                 340                 345

Lys His Thr Met Ile Leu Met Tyr Asn Trp Phe Thr Ser Ser Val
                350                 355                 360

Leu Tyr Gln Gly Leu Ile Met His Met Gly Leu Ala Gly Asp Asn
                365                 370                 375
```

```
 Ile Tyr Leu Asp Phe Phe Tyr Ser Ala Leu Val Glu Phe Pro Ala
                 380                 385                 390

 Ala Phe Met Ile Ile Leu Thr Ile Asp Arg Ile Gly Arg Arg Tyr
                 395                 400                 405

 Pro Trp Ala Ala Ser Asn Met Val Ala Gly Ala Ala Cys Leu Ala
                 410                 415                 420

 Ser Val Phe Ile Pro Gly Asp Leu Gln Trp Leu Lys Ile Ile Ile
                 425                 430                 435

 Ser Cys Leu Gly Arg Met Gly Ile Thr Met Ala Tyr Glu Ile Val
                 440                 445                 450

 Cys Leu Val Asn Ala Glu Leu Tyr Pro Thr Phe Ile Arg Asn Leu
                 455                 460                 465

 Gly Val His Ile Cys Ser Ser Met Cys Asp Ile Gly Gly Ile Ile
                 470                 475                 480

 Thr Pro Phe Leu Val Tyr Arg Leu Thr Asn Ile Trp Leu Glu Leu
                 485                 490                 495

 Pro Leu Met Val Phe Gly Val Leu Gly Leu Val Ala Gly Gly Leu
                 500                 505                 510

 Val Leu Leu Leu Pro Glu Thr Lys Gly Lys Ala Leu Pro Glu Thr
                 515                 520                 525

 Ile Glu Glu Ala Glu Asn Met Gln Arg Pro Arg Lys Asn Lys Glu
                 530                 535                 540

 Lys Met Ile Tyr Leu Gln Val Gln Lys Leu Asp Ile Pro Leu Asn
                 545                 550                 555

<210> 106
<211> 690
<212> PRT
<213> Homo sapiens

<400> 106
 Met Ala Pro Trp Pro Glu Leu Gly Asp Ala Gln Pro Asn Pro Asp
   1               5                  10                  15

 Lys Tyr Leu Glu Gly Ala Ala Gly Gln Gln Pro Thr Ala Pro Asp
                  20                  25                  30

 Lys Ser Lys Glu Thr Asn Lys Thr Asp Asn Thr Glu Ala Pro Val
                  35                  40                  45

 Thr Lys Ile Glu Leu Leu Pro Ser Tyr Ser Thr Ala Thr Leu Ile
                  50                  55                  60

 Asp Glu Pro Thr Glu Val Asp Asp Pro Trp Asn Leu Pro Thr Leu
                  65                  70                  75

 Gln Asp Ser Gly Ile Lys Trp Ser Glu Arg Asp Thr Lys Gly Lys
                  80                  85                  90
```

```
Ile Leu Cys Phe Phe Gln Gly Ile Gly Arg Leu Ile Leu Leu Leu
                 95                  100                 105

Gly Phe Leu Tyr Phe Phe Val Cys Ser Leu Asp Ile Leu Ser Ser
                110                 115                 120

Ala Phe Gln Leu Val Gly Gly Lys Met Ala Gly Gln Phe Phe Ser
                125                 130                 135

Asn Ser Ser Ile Met Ser Asn Pro Leu Leu Gly Leu Val Ile Gly
                140                 145                 150

Val Leu Val Thr Val Leu Val Gln Ser Ser Ser Thr Ser Thr Ser
                155                 160                 165

Ile Val Val Ser Met Val Ser Ser Ser Leu Leu Thr Val Arg Ala
                170                 175                 180

Ala Ile Pro Ile Ile Met Gly Ala Asn Ile Gly Thr Ser Ile Thr
                185                 190                 195

Asn Thr Ile Val Ala Leu Met Gln Val Gly Asp Arg Ser Glu Phe
                200                 205                 210

Arg Arg Ala Phe Ala Gly Ala Thr Val His Asp Phe Phe Asn Trp
                215                 220                 225

Leu Ser Val Leu Val Leu Leu Pro Val Glu Val Ala Thr His Tyr
                230                 235                 240

Leu Glu Ile Ile Thr Gln Leu Ile Val Glu Ser Phe His Phe Lys
                245                 250                 255

Asn Gly Glu Asp Ala Pro Asp Leu Leu Lys Val Ile Thr Lys Pro
                260                 265                 270

Phe Thr Lys Leu Ile Val Gln Leu Asp Lys Lys Val Ile Ser Gln
                275                 280                 285

Ile Ala Met Asn Asp Glu Lys Ala Lys Asn Lys Ser Leu Val Lys
                290                 295                 300

Ile Trp Cys Lys Thr Phe Thr Asn Lys Thr Gln Ile Asn Val Thr
                305                 310                 315

Val Pro Ser Thr Ala Asn Cys Thr Ser Pro Ser Leu Cys Trp Thr
                320                 325                 330

Asp Gly Ile Gln Asn Trp Thr Met Lys Asn Val Thr Tyr Lys Glu
                335                 340                 345

Asn Ile Ala Lys Cys Gln His Ile Phe Val Asn Phe His Leu Pro
                350                 355                 360

Asp Leu Ala Val Gly Thr Ile Leu Leu Ile Leu Ser Leu Leu Val
                365                 370                 375

Leu Cys Gly Cys Leu Ile Met Ile Val Lys Ile Leu Gly Ser Val
                380                 385                 390

Leu Lys Gly Gln Val Ala Thr Val Ile Lys Lys Thr Ile Asn Thr
```

395 400 405

Asp Phe Pro Phe Pro Phe Ala Trp Leu Thr Gly Tyr Leu Ala Ile
410 415 420

Leu Val Gly Ala Gly Met Thr Phe Ile Val Gln Ser Ser Ser Val
425 430 435

Phe Thr Ser Ala Leu Thr Pro Leu Ile Gly Ile Gly Val Ile Thr
440 445 450

Ile Glu Arg Ala Tyr Pro Leu Thr Leu Gly Ser Asn Ile Gly Thr
455 460 465

Thr Thr Thr Ala Ile Leu Ala Ala Leu Ala Ser Pro Gly Asn Ala
470 475 480

Leu Arg Ser Ser Leu Gln Ile Ala Leu Cys His Phe Phe Phe Asn
485 490 495

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

Pro Ile Arg Met Ala Lys Gly Leu Gly Asn Ile Ser Ala Lys Tyr
515 520 525

Arg Trp Phe Ala Val Phe Tyr Leu Ile Ile Phe Phe Phe Leu Ile
530 535 540

Pro Leu Thr Val Phe Gly Leu Ser Leu Ala Gly Trp Arg Val Leu
545 550 555

Val Gly Val Gly Val Pro Val Val Phe Ile Ile Ile Leu Val Leu
560 565 570

Cys Leu Arg Leu Leu Gln Ser Arg Cys Pro Arg Val Leu Pro Lys
575 580 585

Lys Leu Gln Asn Trp Asn Phe Leu Pro Leu Trp Met Arg Ser Leu
590 595 600

Lys Pro Trp Asp Ala Val Val Ser Lys Phe Thr Gly Cys Phe Gln
605 610 615

Met Arg Cys Cys Tyr Cys Cys Arg Val Cys Cys Arg Ala Cys Cys
620 625 630

Leu Leu Cys Gly Cys Pro Lys Cys Cys Arg Cys Ser Lys Cys Cys
635 640 645

Glu Asp Leu Glu Glu Ala Gln Glu Gly Gln Asp Val Pro Val Lys
650 655 660

Ala Pro Glu Thr Phe Asp Asn Ile Thr Ile Ser Arg Glu Ala Gln
665 670 675

Gly Glu Val Pro Ala Ser Asp Ser Lys Thr Glu Cys Thr Ala Leu
680 685 690

<210> 107
<211> 690

<212> PRT
<213> Homo sapiens

<400> 107

```
 Met Ala Pro Trp Pro Glu Leu Gly Asp Ala Gln Pro Asn Pro Asp
   1               5                  10                  15

 Lys Tyr Leu Glu Gly Ala Ala Gly Gln Gln Pro Thr Ala Pro Asp
                20                  25                  30

 Lys Ser Lys Glu Thr Asn Lys Thr Asp Asn Thr Glu Ala Pro Val
                35                  40                  45

 Thr Lys Ile Glu Leu Leu Pro Ser Tyr Ser Thr Ala Thr Leu Ile
                50                  55                  60

 Asp Glu Pro Thr Glu Val Asp Asp Pro Trp Asn Leu Pro Thr Leu
                65                  70                  75

 Gln Asp Ser Gly Ile Lys Trp Ser Glu Arg Asp Thr Lys Gly Lys
                80                  85                  90

 Ile Leu Cys Phe Phe Gln Gly Ile Gly Arg Leu Ile Leu Leu Leu
                95                 100                 105

 Gly Phe Leu Tyr Phe Phe Val Cys Ser Leu Asp Ile Leu Ser Ser
               110                 115                 120

 Ala Phe Gln Leu Val Gly Gly Lys Met Ala Gly Gln Phe Phe Ser
               125                 130                 135

 Asn Ser Ser Ile Met Ser Asn Pro Leu Leu Gly Leu Val Ile Gly
               140                 145                 150

 Val Leu Val Thr Val Leu Val Gln Ser Ser Ser Thr Ser Thr Ser
               155                 160                 165

 Ile Val Val Ser Met Val Ser Ser Ser Leu Leu Thr Val Arg Ala
               170                 175                 180

 Ala Ile Pro Ile Ile Met Gly Ala Asn Ile Gly Thr Ser Ile Thr
               185                 190                 195

 Asn Thr Ile Val Ala Leu Met Gln Val Gly Asp Arg Ser Glu Phe
               200                 205                 210

 Arg Arg Ala Phe Ala Gly Ala Thr Val His Asp Phe Phe Asn Trp
               215                 220                 225

 Leu Ser Val Leu Val Leu Leu Pro Val Glu Val Ala Thr His Tyr
               230                 235                 240

 Leu Glu Ile Ile Thr Gln Leu Ile Val Glu Ser Phe His Phe Lys
               245                 250                 255

 Asn Gly Glu Asp Ala Pro Asp Leu Leu Lys Val Ile Thr Lys Pro
               260                 265                 270

 Phe Thr Lys Leu Ile Val Gln Leu Asp Lys Lys Val Ile Ser Gln
               275                 280                 285
```

```
Ile Ala Met Asn Asp Glu Lys Ala Lys Asn Lys Ser Leu Val Lys
                290                 295                 300

Ile Trp Cys Lys Thr Phe Thr Asn Lys Thr Gln Ile Asn Val Thr
                305                 310                 315

Val Pro Ser Thr Ala Asn Cys Thr Ser Pro Ser Leu Cys Trp Thr
                320                 325                 330

Asp Gly Ile Gln Asn Trp Thr Met Lys Asn Val Thr Tyr Lys Glu
                335                 340                 345

Asn Ile Ala Lys Cys Gln His Ile Phe Val Asn Phe His Leu Pro
                350                 355                 360

Asp Leu Ala Val Gly Thr Ile Leu Leu Ile Leu Ser Leu Leu Val
                365                 370                 375

Leu Cys Gly Cys Leu Ile Met Ile Val Lys Ile Leu Gly Ser Val
                380                 385                 390

Leu Lys Gly Gln Val Ala Thr Val Ile Lys Lys Thr Ile Asn Thr
                395                 400                 405

Asp Phe Pro Phe Pro Phe Ala Trp Leu Thr Gly Tyr Leu Ala Ile
                410                 415                 420

Leu Val Gly Ala Gly Met Thr Phe Ile Val Gln Ser Ser Ser Val
                425                 430                 435

Phe Thr Ser Ala Leu Thr Pro Leu Ile Gly Ile Gly Val Ile Thr
                440                 445                 450

Ile Glu Arg Ala Tyr Pro Leu Thr Leu Gly Ser Asn Ile Gly Thr
                455                 460                 465

Thr Thr Thr Ala Ile Leu Ala Ala Leu Ala Ser Pro Gly Asn Ala
                470                 475                 480

Leu Arg Ser Ser Leu Gln Ile Ala Leu Cys His Phe Phe Phe Asn
                485                 490                 495

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

Pro Ile Arg Met Ala Lys Gly Leu Gly Asn Ile Ser Ala Lys Tyr
                515                 520                 525

Arg Trp Phe Ala Val Phe Tyr Leu Ile Ile Phe Phe Phe Leu Ile
                530                 535                 540

Pro Leu Thr Val Phe Gly Leu Ser Leu Ala Gly Trp Arg Val Leu
                545                 550                 555

Val Gly Val Gly Val Pro Val Val Phe Ile Ile Ile Leu Val Leu
                560                 565                 570

Cys Leu Arg Leu Leu Gln Ser Arg Cys Pro Arg Val Leu Pro Lys
                575                 580                 585

Lys Leu Gln Asn Trp Asn Phe Leu Pro Leu Trp Met Arg Ser Leu
```

```
                590                 595                 600

 Lys Pro Trp Asp Ala Val Val Ser Lys Phe Thr Gly Cys Phe Gln
                605                 610                 615

 Met Arg Cys Cys Cys Cys Cys Arg Val Cys Cys Arg Ala Cys Cys
                620                 625                 630

 Leu Leu Cys Gly Cys Pro Lys Cys Cys Arg Cys Ser Lys Cys Cys
                635                 640                 645

 Glu Asp Leu Glu Glu Ala Gln Glu Gly Gln Asp Val Pro Val Lys
                650                 655                 660

 Ala Pro Glu Thr Phe Asp Asn Ile Thr Ile Ser Arg Glu Ala Gln
                665                 670                 675

 Gly Glu Val Pro Ala Ser Asp Ser Lys Thr Glu Cys Thr Ala Leu
                680                 685                 690

<210> 108
<211> 411
<212> PRT
<213> Homo sapiens

<400> 108
 Met Trp Ser Gly Trp Trp Leu Trp Pro Leu Val Ala Val Cys Thr
   1               5                  10                  15

 Ala Asp Phe Phe Arg Asp Glu Ala Glu Arg Ile Met Arg Asp Ser
                20                  25                  30

 Pro Val Ile Asp Gly His Asn Asp Leu Pro Trp Gln Leu Leu Asp
                35                  40                  45

 Met Phe Asn Asn Arg Leu Gln Asp Glu Arg Ala Asn Leu Thr Thr
                50                  55                  60

 Leu Ala Gly Thr His Thr Asn Ile Pro Lys Leu Arg Ala Gly Phe
                65                  70                  75

 Val Gly Gly Gln Phe Trp Ser Val Tyr Thr Pro Cys Asp Thr Gln
                80                  85                  90

 Asn Lys Asp Ala Val Arg Arg Thr Leu Glu Gln Met Asp Val Val
                95                 100                 105

 His Arg Met Cys Arg Met Tyr Pro Glu Thr Phe Leu Tyr Val Thr
               110                 115                 120

 Ser Ser Ala Gly Ile Arg Gln Ala Phe Arg Glu Gly Lys Val Ala
               125                 130                 135

 Ser Leu Ile Gly Val Glu Gly Gly His Ser Ile Asp Ser Ser Leu
               140                 145                 150

 Gly Val Leu Arg Ala Leu Tyr Gln Leu Gly Met Arg Tyr Leu Thr
               155                 160                 165

 Leu Thr His Ser Cys Asn Thr Pro Trp Ala Asp Asn Trp Leu Val
               170                 175                 180
```

```
 Asp Thr Gly Asp Ser Glu Pro Gln Ser Gln Gly Leu Ser Pro Phe
                 185                 190                 195

 Gly Gln Arg Val Val Lys Glu Leu Asn Arg Leu Gly Val Leu Ile
                 200                 205                 210

 Asp Leu Ala His Val Ser Val Ala Thr Met Lys Ala Thr Leu Gln
                 215                 220                 225

 Leu Ser Arg Ala Pro Val Ile Phe Ser His Ser Ser Ala Tyr Ser
                 230                 235                 240

 Val Cys Ala Ser Arg Arg Asn Val Pro Asp Asp Val Leu Arg Leu
                 245                 250                 255

 Val Lys Gln Thr Asp Ser Leu Val Met Val Asn Phe Tyr Asn Asn
                 260                 265                 270

 Tyr Ile Ser Cys Thr Asn Lys Ala Asn Leu Ser Gln Val Ala Asp
                 275                 280                 285

 His Leu Asp His Ile Lys Glu Val Ala Gly Ala Arg Ala Val Gly
                 290                 295                 300

 Phe Gly Gly Asp Phe Asp Gly Val Pro Arg Val Pro Glu Gly Leu
                 305                 310                 315

 Glu Asp Val Ser Lys Tyr Pro Asp Leu Ile Ala Glu Leu Leu Arg
                 320                 325                 330

 Arg Asn Trp Thr Glu Ala Glu Val Lys Gly Ala Leu Ala Asp Asn
                 335                 340                 345

 Leu Leu Arg Val Phe Gln Ala Val Glu Gln Ala Ser Asn Leu Thr
                 350                 355                 360

 Gln Ala Pro Glu Glu Glu Pro Ile Pro Leu Asp Gln Leu Gly Gly
                 365                 370                 375

 Ser Cys Arg Thr His Tyr Gly Tyr Ser Ser Gly Ala Ser Ser Leu
                 380                 385                 390

 His Arg His Trp Gly Leu Leu Leu Ala Ser Leu Ala Pro Leu Val
                 395                 400                 405

 Leu Cys Leu Ser Leu Leu
                 410

<210> 109
<211> 320
<212> PRT
<213> Homo sapiens

<400> 109
 Met Ser Ser Cys Asn Phe Thr His Ala Thr Phe Val Leu Ile Gly
   1               5                  10                  15

 Ile Pro Gly Leu Glu Lys Ala His Phe Trp Val Gly Phe Pro Leu
                  20                  25                  30
```

```
Leu Ser Met Tyr Val Val Ala Met Phe Gly Asn Cys Ile Val Val
                 35                  40                  45

Phe Ile Val Arg Thr Glu Arg Ser Leu His Ala Pro Met Tyr Leu
                 50                  55                  60

Phe Leu Cys Met Leu Ala Ala Ile Asp Leu Ala Leu Ser Thr Ser
                 65                  70                  75

Thr Met Pro Lys Ile Leu Ala Leu Phe Trp Phe Asp Ser Arg Glu
                 80                  85                  90

Ile Ser Phe Glu Ala Cys Leu Thr Gln Met Phe Phe Ile His Ala
                 95                 100                 105

Leu Ser Ala Ile Glu Ser Thr Ile Leu Leu Ala Met Ala Phe Asp
                110                 115                 120

Arg Tyr Val Ala Ile Cys His Pro Leu Arg His Ala Ala Val Leu
                125                 130                 135

Asn Asn Thr Val Thr Ala Gln Ile Gly Ile Val Ala Val Val Arg
                140                 145                 150

Gly Ser Leu Phe Phe Phe Pro Leu Pro Leu Leu Ile Lys Arg Leu
                155                 160                 165

Ala Phe Cys His Ser Asn Val Leu Ser His Ser Tyr Cys Val His
                170                 175                 180

Gln Asp Val Met Lys Leu Ala Tyr Ala Asp Thr Leu Pro Asn Val
                185                 190                 195

Val Tyr Gly Leu Thr Ala Ile Leu Leu Val Met Gly Val Asp Val
                200                 205                 210

Met Phe Ile Ser Leu Ser Tyr Phe Leu Ile Ile Arg Thr Val Leu
                215                 220                 225

Gln Leu Pro Ser Lys Ser Glu Arg Ala Lys Ala Phe Gly Thr Cys
                230                 235                 240

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

Gly Leu Ser Val Val His Arg Phe Gly Asn Ser Leu His Pro Ile
                260                 265                 270

Val Arg Val Val Met Gly Asp Ile Tyr Leu Leu Leu Pro Pro Val
                275                 280                 285

Ile Asn Pro Ile Ile Tyr Gly Ala Lys Thr Lys Gln Ile Arg Thr
                290                 295                 300

Arg Val Leu Ala Met Phe Lys Ile Ser Cys Asp Lys Asp Leu Gln
                305                 310                 315

Ala Val Gly Gly Lys
                320
```

<210> 110

<211> 2314
<212> PRT
<213> Homo sapiens

<400> 110
Met Arg Ile Leu Lys Arg Phe Leu Ala Cys Ile Gln Leu Leu Cys
1 5 10 15

Val Cys Arg Leu Asp Trp Ala Asn Gly Tyr Tyr Arg Gln Gln Arg
20 25 30

Lys Leu Val Glu Glu Ile Gly Trp Ser Tyr Thr Gly Ala Leu Asn
35 40 45

Gln Lys Asn Trp Gly Lys Lys Tyr Pro Thr Cys Asn Ser Pro Lys
50 55 60

Gln Ser Pro Ile Asn Ile Asp Glu Asp Leu Thr Gln Val Asn Val
65 70 75

Asn Leu Lys Lys Leu Lys Phe Gln Gly Trp Asp Lys Thr Ser Leu
80 85 90

Glu Asn Thr Phe Ile His Asn Thr Gly Lys Thr Val Glu Ile Asn
95 100 105

Leu Thr Asn Asp Tyr Arg Val Ser Gly Gly Val Ser Glu Met Val
110 115 120

Phe Lys Ala Ser Lys Ile Thr Phe His Trp Gly Lys Cys Asn Met
125 130 135

Ser Ser Asp Gly Ser Glu His Ser Leu Glu Gly Gln Lys Phe Pro
140 145 150

Leu Glu Met Gln Ile Tyr Cys Phe Asp Ala Asp Arg Phe Ser Ser
155 160 165

Phe Glu Glu Ala Val Lys Gly Lys Gly Lys Leu Arg Ala Leu Ser
170 175 180

Ile Leu Phe Glu Val Gly Thr Glu Glu Asn Leu Asp Phe Lys Ala
185 190 195

Ile Ile Asp Gly Val Glu Ser Val Ser Arg Phe Gly Lys Gln Ala
200 205 210

Ala Leu Asp Pro Phe Ile Leu Leu Asn Leu Leu Pro Asn Ser Thr
215 220 225

Asp Lys Tyr Tyr Ile Tyr Asn Gly Ser Leu Thr Ser Pro Pro Cys
230 235 240

Thr Asp Thr Val Asp Trp Ile Val Phe Lys Asp Thr Val Ser Ile
245 250 255

Ser Glu Ser Gln Leu Ala Val Phe Cys Glu Val Leu Thr Met Gln
260 265 270

Gln Ser Gly Tyr Val Met Leu Met Asp Tyr Leu Gln Asn Asn Phe
275 280 285

```
Arg Glu Gln Gln Tyr Lys Phe Ser Arg Gln Val Phe Ser Ser Tyr
                290                 295                 300

Thr Gly Lys Glu Glu Ile His Glu Ala Val Cys Ser Ser Glu Pro
                305                 310                 315

Glu Asn Val Gln Ala Asp Pro Glu Asn Tyr Thr Ser Leu Leu Val
                320                 325                 330

Thr Trp Glu Arg Pro Arg Val Val Tyr Asp Thr Met Ile Glu Lys
                335                 340                 345

Phe Ala Val Leu Tyr Gln Gln Leu Asp Gly Glu Asp Gln Thr Lys
                350                 355                 360

His Glu Phe Leu Thr Asp Gly Tyr Gln Asp Leu Gly Ala Ile Leu
                365                 370                 375

Asn Asn Leu Leu Pro Asn Met Ser Tyr Val Leu Gln Ile Val Ala
                380                 385                 390

Ile Cys Thr Asn Gly Leu Tyr Gly Lys Tyr Ser Asp Gln Leu Ile
                395                 400                 405

Val Asp Met Pro Thr Asp Asn Pro Glu Leu Asp Leu Phe Pro Glu
                410                 415                 420

Leu Ile Gly Thr Glu Glu Ile Ile Lys Glu Glu Glu Glu Gly Lys
                425                 430                 435

Asp Ile Glu Glu Gly Ala Ile Val Asn Pro Gly Arg Asp Ser Ala
                440                 445                 450

Thr Asn Gln Ile Arg Lys Lys Glu Pro Gln Ile Ser Thr Thr Thr
                455                 460                 465

His Tyr Asn Arg Ile Gly Thr Lys Tyr Asn Glu Ala Lys Thr Asn
                470                 475                 480

Arg Ser Pro Thr Arg Gly Ser Glu Phe Ser Gly Lys Gly Asp Val
                485                 490                 495

Pro Asn Thr Ser Leu Asn Ser Thr Ser Gln Pro Val Thr Lys Leu
                500                 505                 510

Ala Thr Glu Lys Asp Ile Ser Leu Thr Ser Gln Thr Val Thr Glu
                515                 520                 525

Leu Pro Pro His Thr Val Glu Gly Thr Ser Ala Ser Leu Asn Asp
                530                 535                 540

Gly Ser Lys Thr Val Leu Arg Ser Pro His Met Asn Leu Ser Gly
                545                 550                 555

Thr Ala Glu Ser Leu Asn Thr Val Ser Ile Thr Glu Tyr Glu Glu
                560                 565                 570

Glu Ser Leu Leu Thr Ser Phe Lys Leu Asp Thr Gly Ala Glu Asp
                575                 580                 585
```

```
Ser Ser Gly Ser Ser Pro Ala Thr Ser Ala Ile Pro Phe Ile Ser
                590                 595                 600

Glu Asn Ile Ser Gln Gly Tyr Ile Phe Ser Ser Glu Asn Pro Glu
                605                 610                 615

Thr Ile Thr Tyr Asp Val Leu Ile Pro Glu Ser Ala Arg Asn Ala
                620                 625                 630

Ser Glu Asp Ser Thr Ser Ser Gly Ser Glu Glu Ser Leu Lys Asp
                635                 640                 645

Pro Ser Met Glu Gly Asn Val Trp Phe Pro Ser Ser Thr Asp Ile
                650                 655                 660

Thr Ala Gln Pro Asp Val Gly Ser Gly Arg Glu Ser Phe Leu Gln
                665                 670                 675

Thr Asn Tyr Thr Glu Ile Arg Val Asp Glu Ser Glu Lys Thr Thr
                680                 685                 690

Lys Ser Phe Ser Ala Gly Pro Val Met Ser Gln Gly Pro Ser Val
                695                 700                 705

Thr Asp Leu Glu Met Pro His Tyr Ser Thr Phe Ala Tyr Phe Pro
                710                 715                 720

Thr Glu Val Thr Pro His Ala Phe Thr Pro Ser Ser Arg Gln Gln
                725                 730                 735

Asp Leu Val Ser Thr Val Asn Val Val Tyr Ser Gln Thr Thr Gln
                740                 745                 750

Pro Val Tyr Asn Gly Glu Thr Pro Leu Gln Pro Ser Tyr Ser Ser
                755                 760                 765

Glu Val Phe Pro Leu Val Thr Pro Leu Leu Leu Asp Asn Gln Ile
                770                 775                 780

Leu Asn Thr Thr Pro Ala Ala Ser Ser Ser Asp Ser Ala Leu His
                785                 790                 795

Ala Thr Pro Val Phe Pro Ser Val Asp Val Ser Phe Glu Ser Ile
                800                 805                 810

Leu Ser Ser Tyr Asp Gly Ala Pro Leu Leu Pro Phe Ser Ser Ala
                815                 820                 825

Ser Phe Ser Ser Glu Leu Phe Arg His Leu His Thr Val Ser Gln
                830                 835                 840

Ile Leu Pro Gln Val Thr Ser Ala Thr Glu Ser Asp Lys Val Pro
                845                 850                 855

Leu His Ala Ser Leu Pro Val Ala Gly Gly Asp Leu Leu Leu Glu
                860                 865                 870

Pro Ser Leu Ala Gln Tyr Ser Asp Val Leu Ser Thr Thr His Ala
                875                 880                 885

Ala Ser Glu Thr Leu Glu Phe Gly Ser Glu Ser Gly Val Leu Tyr
```

```
                890                 895                 900

Lys Thr Leu Met Phe Ser Gln Val Glu Pro Pro Ser Ser Asp Ala
                905                 910                 915

Met Met His Ala Arg Ser Ser Gly Pro Glu Pro Ser Tyr Ala Leu
                920                 925                 930

Ser Asp Asn Glu Gly Ser Gln His Ile Phe Thr Val Ser Tyr Ser
                935                 940                 945

Ser Ala Ile Pro Val His Asp Ser Val Gly Val Thr Tyr Gln Gly
                950                 955                 960

Ser Leu Phe Ser Gly Pro Ser His Ile Pro Ile Pro Lys Ser Ser
                965                 970                 975

Leu Ile Thr Pro Thr Ala Ser Leu Leu Gln Pro Thr His Ala Leu
                980                 985                 990

Ser Gly Asp Gly Glu Trp Ser Gly Ala Ser Ser Asp Ser Glu Phe
                995                1000                1005

Leu Leu Pro Asp Thr Asp Gly Leu Thr Ala Leu Asn Ile Ser Ser
               1010                1015                1020

Pro Val Ser Val Ala Glu Phe Thr Tyr Thr Thr Ser Val Phe Gly
               1025                1030                1035

Asp Asp Asn Lys Ala Leu Ser Lys Ser Glu Ile Ile Tyr Gly Asn
               1040                1045                1050

Glu Thr Glu Leu Gln Ile Pro Ser Phe Asn Glu Met Val Tyr Pro
               1055                1060                1065

Ser Glu Ser Thr Val Met Pro Asn Met Tyr Asp Asn Val Asn Lys
               1070                1075                1080

Leu Asn Ala Ser Leu Gln Glu Thr Ser Val Ser Ile Ser Ser Thr
               1085                1090                1095

Lys Gly Met Phe Pro Gly Ser Leu Ala His Thr Thr Thr Lys Val
               1100                1105                1110

Phe Asp His Glu Ile Ser Gln Val Pro Glu Asn Asn Phe Ser Val
               1115                1120                1125

Gln Pro Thr His Thr Val Ser Gln Ala Ser Gly Asp Thr Ser Leu
               1130                1135                1140

Lys Pro Val Leu Ser Ala Asn Ser Glu Pro Ala Ser Ser Asp Pro
               1145                1150                1155

Ala Ser Ser Glu Met Leu Ser Pro Ser Thr Gln Leu Leu Phe Tyr
               1160                1165                1170

Glu Thr Ser Ala Ser Phe Ser Thr Glu Val Leu Leu Gln Pro Ser
               1175                1180                1185

Phe Gln Ala Ser Asp Val Asp Thr Leu Leu Lys Thr Val Leu Pro
               1190                1195                1200
```

```
Ala Val Pro Ser Asp Pro Ile Leu Val Glu Thr Pro Lys Val Asp
                1205                1210                1215

Lys Ile Ser Ser Thr Met Leu His Leu Ile Val Ser Asn Ser Ala
                1220                1225                1230

Ser Ser Glu Asn Met Leu His Ser Thr Ser Val Pro Val Phe Asp
                1235                1240                1245

Val Ser Pro Thr Ser His Met His Ser Ala Ser Leu Gln Gly Leu
                1250                1255                1260

Thr Ile Ser Tyr Ala Ser Glu Lys Tyr Glu Pro Val Leu Leu Lys
                1265                1270                1275

Ser Glu Ser Ser His Gln Val Val Pro Ser Leu Tyr Ser Asn Asp
                1280                1285                1290

Glu Leu Phe Gln Thr Ala Asn Leu Glu Ile Asn Gln Ala His Pro
                1295                1300                1305

Pro Lys Gly Arg His Val Phe Ala Thr Pro Val Leu Ser Ile Asp
                1310                1315                1320

Glu Pro Leu Asn Thr Leu Ile Asn Lys Leu Ile His Ser Asp Glu
                1325                1330                1335

Ile Leu Thr Ser Thr Lys Ser Ser Val Thr Gly Lys Val Phe Ala
                1340                1345                1350

Gly Ile Pro Thr Val Ala Ser Asp Thr Phe Val Ser Thr Asp His
                1355                1360                1365

Ser Val Pro Ile Gly Asn Gly His Val Ala Ile Thr Ala Val Ser
                1370                1375                1380

Pro His Arg Asp Gly Ser Val Thr Ser Thr Lys Leu Leu Phe Pro
                1385                1390                1395

Ser Lys Ala Thr Ser Glu Leu Ser His Ser Ala Lys Ser Asp Ala
                1400                1405                1410

Gly Leu Val Gly Gly Gly Glu Asp Gly Asp Thr Asp Asp Asp Gly
                1415                1420                1425

Asp Asp Asp Asp Asp Arg Asp Ser Asp Gly Leu Ser Ile His Lys
                1430                1435                1440

Cys Met Ser Cys Ser Ser Tyr Arg Glu Ser Gln Glu Lys Val Met
                1445                1450                1455

Asn Asp Ser Asp Thr His Glu Asn Ser Leu Met Asp Gln Asn Asn
                1460                1465                1470

Pro Ile Ser Tyr Ser Leu Ser Glu Asn Ser Glu Glu Asp Asn Arg
                1475                1480                1485

Val Thr Ser Val Ser Ser Asp Ser Gln Thr Gly Met Asp Arg Ser
                1490                1495                1500
```

```
Pro Gly Lys Ser Pro Ser Ala Asn Gly Leu Ser Gln Lys His Asn
                1505                1510                1515

Asp Gly Lys Glu Glu Asn Asp Ile Gln Thr Gly Ser Ala Leu Leu
                1520                1525                1530

Pro Leu Ser Pro Glu Ser Lys Ala Trp Ala Val Leu Thr Ser Asp
                1535                1540                1545

Glu Glu Ser Gly Ser Gly Gln Gly Thr Ser Asp Ser Leu Asn Glu
                1550                1555                1560

Asn Glu Thr Ser Thr Asp Phe Ser Phe Ala Asp Thr Asn Glu Lys
                1565                1570                1575

Asp Ala Asp Gly Ile Leu Ala Ala Gly Asp Ser Glu Ile Thr Pro
                1580                1585                1590

Gly Phe Pro Gln Ser Pro Thr Ser Ser Val Thr Ser Glu Asn Ser
                1595                1600                1605

Glu Val Phe His Val Ser Glu Ala Glu Ala Ser Asn Ser Ser His
                1610                1615                1620

Glu Ser Arg Ile Gly Leu Ala Glu Gly Leu Glu Ser Glu Lys Lys
                1625                1630                1635

Ala Val Ile Pro Leu Val Ile Val Ser Ala Leu Thr Phe Ile Cys
                1640                1645                1650

Leu Val Val Leu Val Gly Ile Leu Ile Tyr Trp Arg Lys Cys Phe
                1655                1660                1665

Gln Thr Ala His Phe Tyr Leu Glu Asp Ser Thr Ser Pro Arg Val
                1670                1675                1680

Ile Ser Thr Pro Pro Thr Pro Ile Phe Pro Ile Ser Asp Asp Val
                1685                1690                1695

Gly Ala Ile Pro Ile Lys His Phe Pro Lys His Val Ala Asp Leu
                1700                1705                1710

His Ala Ser Ser Gly Phe Thr Glu Glu Phe Glu Thr Leu Lys Glu
                1715                1720                1725

Phe Tyr Gln Glu Val Gln Ser Cys Thr Val Asp Leu Gly Ile Thr
                1730                1735                1740

Ala Asp Ser Ser Asn His Pro Asp Asn Lys His Lys Asn Arg Tyr
                1745                1750                1755

Ile Asn Ile Val Ala Tyr Asp His Ser Arg Val Lys Leu Ala Gln
                1760                1765                1770

Leu Ala Glu Lys Asp Gly Lys Leu Thr Asp Tyr Ile Asn Ala Asn
                1775                1780                1785

Tyr Val Asp Gly Tyr Asn Arg Pro Lys Ala Tyr Ile Ala Ala Gln
                1790                1795                1800

Gly Pro Leu Lys Ser Thr Ala Glu Asp Phe Trp Arg Met Ile Trp
```

```
                1805                1810                1815

Glu His Asn Val Glu Val Ile Val Met Ile Thr Asn Leu Val Glu
                1820                1825                1830

Lys Gly Arg Arg Lys Cys Asp Gln Tyr Trp Pro Ala Asp Gly Ser
                1835                1840                1845

Glu Glu Tyr Gly Asn Phe Leu Val Thr Gln Lys Ser Val Gln Val
                1850                1855                1860

Leu Ala Tyr Tyr Thr Val Arg Asn Phe Thr Leu Arg Asn Thr Lys
                1865                1870                1875

Ile Lys Lys Gly Ser Gln Lys Gly Arg Pro Ser Gly Arg Val Val
                1880                1885                1890

Thr Gln Tyr His Tyr Thr Gln Trp Pro Asp Met Gly Val Pro Glu
                1895                1900                1905

Tyr Ser Leu Pro Val Leu Thr Phe Val Arg Lys Ala Ala Tyr Ala
                1910                1915                1920

Lys Arg His Ala Val Gly Pro Val Val Val His Cys Ser Ala Gly
                1925                1930                1935

Val Gly Arg Thr Gly Thr Tyr Ile Val Leu Asp Ser Met Leu Gln
                1940                1945                1950

Gln Ile Gln His Glu Gly Thr Val Asn Ile Phe Gly Phe Leu Lys
                1955                1960                1965

His Ile Arg Ser Gln Arg Asn Tyr Leu Val Gln Thr Glu Glu Gln
                1970                1975                1980

Tyr Val Phe Ile His Asp Thr Leu Val Glu Ala Ile Leu Ser Lys
                1985                1990                1995

Glu Thr Glu Val Leu Asp Ser His Ile His Ala Tyr Val Asn Ala
                2000                2005                2010

Leu Leu Ile Pro Gly Pro Ala Gly Lys Thr Lys Leu Glu Lys Gln
                2015                2020                2025

Phe Gln Leu Leu Ser Gln Ser Asn Ile Gln Gln Ser Asp Tyr Ser
                2030                2035                2040

Ala Ala Leu Lys Gln Cys Asn Arg Glu Lys Asn Arg Thr Ser Ser
                2045                2050                2055

Ile Ile Pro Val Glu Arg Ser Arg Val Gly Ile Ser Ser Leu Ser
                2060                2065                2070

Gly Glu Gly Thr Asp Tyr Ile Asn Ala Ser Tyr Ile Met Gly Tyr
                2075                2080                2085

Tyr Gln Ser Asn Glu Phe Ile Ile Thr Gln His Pro Leu Leu His
                2090                2095                2100

Thr Ile Lys Asp Phe Trp Arg Met Ile Trp Asp His Asn Ala Gln
                2105                2110                2115
```

```
 Leu Val Val Met Ile Pro Asp Gly Gln Asn Met Ala Glu Asp Glu
                2120                2125                2130

 Phe Val Tyr Trp Pro Asn Lys Asp Glu Pro Ile Asn Cys Glu Ser
                2135                2140                2145

 Phe Lys Val Thr Leu Met Ala Glu Glu His Lys Cys Leu Ser Asn
                2150                2155                2160

 Glu Glu Lys Leu Ile Ile Gln Asp Phe Ile Leu Glu Ala Thr Gln
                2165                2170                2175

 Asp Asp Tyr Val Leu Glu Val Arg His Phe Gln Cys Pro Lys Trp
                2180                2185                2190

 Pro Asn Pro Asp Ser Pro Ile Ser Lys Thr Phe Glu Leu Ile Ser
                2195                2200                2205

 Val Ile Lys Glu Glu Ala Ala Asn Arg Asp Gly Pro Met Ile Val
                2210                2215                2220

 His Asp Glu His Gly Gly Val Thr Ala Gly Thr Phe Cys Ala Leu
                2225                2230                2235

 Thr Thr Leu Met His Gln Leu Glu Lys Glu Asn Ser Val Asp Val
                2240                2245                2250

 Tyr Gln Val Ala Lys Met Ile Asn Leu Met Arg Pro Gly Val Phe
                2255                2260                2265

 Ala Asp Ile Glu Gln Tyr Gln Phe Leu Tyr Lys Val Ile Leu Ser
                2270                2275                2280

 Leu Val Ser Thr Arg Gln Glu Glu Asn Pro Ser Thr Ser Leu Asp
                2285                2290                2295

 Ser Asn Gly Ala Ala Leu Pro Asp Gly Asn Ile Ala Glu Ser Leu
                2300                2305                2310

 Glu Ser Leu Val


<210> 111
<211> 774
<212> PRT
<213> Homo sapiens

<400> 111
 Met Cys Ala Arg Met Ala Gly Arg Thr Arg Ala Ala Pro Arg Gly
   1               5                  10                  15

 Pro Tyr Gly Pro Trp Leu Cys Leu Leu Val Ala Leu Ala Leu Asp
                20                  25                  30

 Val Val Arg Val Asp Cys Gly Gln Ala Pro Leu Asp Pro Val Tyr
                35                  40                  45

 Leu His Val Thr Ala Ala Arg Pro Ala Gln Pro Thr Leu Trp Thr
                50                  55                  60
```

```
Ala Lys Leu Asp Arg Phe Lys Gly Ser Arg His His Thr Thr Leu
                 65                  70                  75

Ile Thr Cys His Arg Ala Gly Leu Thr Glu Pro Asp Ser Ser Ser
                 80                  85                  90

Pro Leu Glu Leu Ser Glu Phe Leu Trp Val Asp Phe Val Val Glu
                 95                 100                 105

Asn Ser Thr Gly Gly Gly Val Ala Val Thr Arg Pro Val Thr Trp
                110                 115                 120

Gln Leu Glu Tyr Pro Gly Gln Ala Pro Glu Ala Glu Lys Asp Lys
                125                 130                 135

Met Val Trp Glu Ile Leu Val Ser Glu Arg Asp Ile Arg Ala Leu
                140                 145                 150

Ile Pro Leu Ala Lys Ala Glu Glu Leu Val Asn Thr Ala Pro Leu
                155                 160                 165

Thr Gly Val Pro Gln His Val Pro Val Arg Leu Val Thr Val Asp
                170                 175                 180

Gly Gly Gly Ala Leu Val Glu Val Thr Glu His Val Gly Cys Glu
                185                 190                 195

Ser Ala Asn Thr Gln Val Leu Gln Val Ser Glu Ala Cys Asp Ala
                200                 205                 210

Val Phe Val Ala Gly Lys Glu Ser Arg Gly Ala Arg Gly Val Arg
                215                 220                 225

Val Asp Phe Trp Trp Arg Arg Leu Arg Ala Ser Leu Arg Leu Thr
                230                 235                 240

Val Trp Ala Pro Leu Leu Pro Leu Arg Ile Glu Leu Thr Asp Thr
                245                 250                 255

Thr Leu Glu Gln Val Arg Gly Trp Arg Val Pro Gly Pro Ala Glu
                260                 265                 270

Gly Pro Ala Glu Pro Ala Ala Glu Ala Ser Asp Glu Ala Glu Arg
                275                 280                 285

Arg Ala Arg Gly Cys His Leu Gln Tyr Gln Arg Ala Gly Val Arg
                290                 295                 300

Phe Leu Ala Pro Phe Ala Ala His Pro Leu Asp Gly Gly Arg Arg
                305                 310                 315

Leu Thr His Leu Leu Gly Pro Asp Trp Leu Leu Asp Val Ser His
                320                 325                 330

Leu Val Ala Pro His Ala Arg Val Leu Asp Ser Arg Val Ala Ser
                335                 340                 345

Leu Glu Gly Gly Arg Val Val Val Gly Arg Glu Pro Gly Val Thr
                350                 355                 360

Ser Ile Glu Val Arg Ser Pro Leu Ser Asp Ser Ile Leu Gly Glu
```

```
                365                 370                 375

Gln Ala Leu Ala Val Thr Asp Asp Lys Val Ser Val Leu Glu Leu
                380                 385                 390

Arg Val Gln Pro Val Met Gly Ile Ser Leu Thr Leu Ser Arg Gly
                395                 400                 405

Thr Ala His Pro Gly Glu Val Thr Ala Thr Cys Trp Ala Gln Ser
                410                 415                 420

Ala Leu Pro Ala Pro Lys Gln Glu Val Ala Leu Ser Leu Trp Leu
                425                 430                 435

Ser Phe Ser Asp His Thr Val Ala Pro Ala Glu Leu Tyr Asp Arg
                440                 445                 450

Arg Asp Leu Gly Leu Ser Val Ser Ala Glu Glu Pro Gly Ala Ile
                455                 460                 465

Leu Pro Ala Glu Glu Gln Gly Ala Gln Leu Gly Val Val Val Ser
                470                 475                 480

Gly Ala Gly Ala Glu Gly Leu Pro Leu His Val Ala Leu His Pro
                485                 490                 495

Pro Glu Pro Cys Arg Arg Gly Arg His Arg Val Pro Leu Ala Ser
                500                 505                 510

Gly Thr Ala Trp Leu Gly Leu Pro Pro Ala Ser Thr Pro Ala Pro
                515                 520                 525

Ala Leu Pro Ser Ser Pro Ala Trp Ser Pro Pro Ala Thr Glu Ala
                530                 535                 540

Thr Met Gly Gly Lys Arg Gln Val Ala Gly Ser Val Gly Gly Asn
                545                 550                 555

Thr Gly Val Arg Gly Lys Phe Glu Arg Ala Glu Glu Glu Ala Arg
                560                 565                 570

Lys Glu Glu Thr Glu Ala Arg Glu Glu Glu Glu Glu Glu Glu Glu
                575                 580                 585

Glu Met Val Pro Ala Pro Gln His Val Thr Glu Leu Glu Leu Gly
                590                 595                 600

Met Tyr Ala Leu Leu Gly Val Phe Cys Val Ala Ile Phe Ile Phe
                605                 610                 615

Leu Val Asn Gly Val Val Phe Val Leu Arg Tyr Gln Arg Lys Glu
                620                 625                 630

Pro Pro Asp Ser Ala Thr Asp Pro Thr Ser Pro Gln Pro His Asn
                635                 640                 645

Trp Val Trp Leu Gly Thr Asp Gln Glu Glu Leu Ser Arg Gln Leu
                650                 655                 660

Asp Arg Gln Ser Pro Gly Pro Pro Lys Gly Glu Gly Ser Cys Pro
                665                 670                 675
```

```
 Cys Glu Ser Gly Gly Gly Gly Glu Ala Pro Thr Leu Ala Pro Gly
                 680                 685                 690

 Pro Pro Gly Gly Thr Thr Ser Ser Ser Ser Thr Leu Ala Arg Lys
                 695                 700                 705

 Glu Ala Gly Gly Arg Arg Lys Arg Val Glu Phe Val Thr Phe Val
                 710                 715                 720

 Pro Ala Pro Pro Ala Gln Ser Pro Glu Glu Pro Val Gly Ala Pro
                 725                 730                 735

 Ala Val Gln Ser Ile Leu Val Ala Gly Glu Glu Asp Ile Arg Trp
                 740                 745                 750

 Val Cys Glu Asp Met Gly Leu Lys Asp Pro Glu Glu Leu Arg Asn
                 755                 760                 765

 Tyr Met Glu Arg Ile Arg Gly Ser Ser
                 770

<210> 112
<211> 420
<212> PRT
<213> Homo sapiens

<400> 112
 Met Ala Thr Lys Thr Glu Leu Ser Pro Thr Ala Arg Glu Ser Lys
   1               5                  10                  15

 Asn Ala Gln Asp Met Gln Val Asp Glu Thr Leu Ile Pro Arg Lys
                  20                  25                  30

 Gly Pro Ser Leu Cys Ser Ala Arg Tyr Gly Ile Ala Leu Val Leu
                  35                  40                  45

 His Phe Cys Asn Phe Thr Thr Ile Ala Gln Asn Val Ile Met Asn
                  50                  55                  60

 Ile Thr Met Val Ala Met Val Asn Ser Thr Ser Pro Gln Ser Gln
                  65                  70                  75

 Leu Asn Asp Ser Ser Glu Val Leu Pro Val Asp Ser Phe Gly Gly
                  80                  85                  90

 Leu Ser Lys Ala Pro Lys Ser Leu Pro Ala Lys Ser Ser Ile Leu
                  95                 100                 105

 Gly Gly Gln Phe Ala Ile Trp Glu Lys Trp Gly Pro Pro Gln Glu
                 110                 115                 120

 Arg Ser Arg Leu Cys Ser Ile Ala Leu Ser Gly Met Leu Leu Gly
                 125                 130                 135

 Cys Phe Thr Ala Ile Leu Ile Gly Gly Phe Ile Ser Glu Thr Leu
                 140                 145                 150

 Gly Trp Pro Phe Val Phe Tyr Ile Phe Gly Gly Val Gly Cys Val
                 155                 160                 165
```

```
 Cys Cys Leu Leu Trp Phe Val Val Ile Tyr Asp Asp Pro Phe Ser
                 170                 175                 180

 Tyr Pro Trp Ile Ser Thr Ser Glu Lys Glu Tyr Ile Ile Ser Ser
                 185                 190                 195

 Leu Lys Gln Gln Val Gly Ser Ser Lys Gln Pro Leu Pro Ile Lys
                 200                 205                 210

 Ala Met Leu Arg Ser Leu Pro Ile Trp Ser Ile Cys Leu Gly Cys
                 215                 220                 225

 Phe Ser His Gln Trp Leu Val Ser Thr Met Val Val Tyr Ile Pro
                 230                 235                 240

 Thr Tyr Ile Ser Ser Val Tyr His Val Asn Ile Arg Asp Asn Gly
                 245                 250                 255

 Leu Leu Ser Ala Leu Pro Phe Ile Val Ala Trp Val Ile Gly Met
                 260                 265                 270

 Val Gly Gly Tyr Leu Ala Asp Phe Leu Leu Thr Lys Lys Phe Arg
                 275                 280                 285

 Leu Ile Thr Val Arg Lys Ile Ala Thr Ile Leu Gly Ser Leu Pro
                 290                 295                 300

 Ser Ser Ala Leu Ile Val Ser Leu Pro Tyr Leu Asn Ser Gly Tyr
                 305                 310                 315

 Ile Thr Ala Thr Ala Leu Leu Thr Leu Ser Cys Gly Leu Ser Thr
                 320                 325                 330

 Leu Cys Gln Ser Gly Ile Tyr Ile Asn Val Leu Asp Ile Ala Pro
                 335                 340                 345

 Arg Tyr Ser Ser Phe Leu Met Gly Ala Ser Arg Gly Phe Ser Ser
                 350                 355                 360

 Ile Ala Pro Val Ile Val Pro Thr Val Ser Gly Phe Leu Leu Ser
                 365                 370                 375

 Gln Asp Pro Glu Phe Gly Trp Arg Asn Val Phe Phe Leu Leu Phe
                 380                 385                 390

 Ala Val Asn Leu Leu Gly Leu Leu Phe Tyr Leu Ile Phe Gly Glu
                 395                 400                 405

 Ala Asp Val Gln Glu Trp Ala Lys Glu Arg Lys Leu Thr Arg Leu
                 410                 415                 420

<210> 113
<211> 401
<212> PRT
<213> Homo sapiens

<400> 113
 Met Gln Val Asp Glu Thr Leu Ile Pro Arg Lys Val Pro Ser Leu
   1               5                  10                  15

 Cys Ser Ala Arg Tyr Gly Ile Ala Leu Val Leu His Phe Cys Asn
```

```
                 20                  25                  30

Phe Thr Thr Ile Ala Gln Asn Val Ile Met Asn Ile Thr Met Val
                 35                  40                  45

Ala Met Val Asn Ser Thr Ser Pro Gln Ser Gln Leu Asn Asp Ser
                 50                  55                  60

Ser Glu Val Leu Pro Val Asp Ser Phe Gly Gly Leu Ser Lys Ala
                 65                  70                  75

Pro Lys Ser Leu Pro Thr Lys Ser Ser Ile Leu Gly Gly Gln Phe
                 80                  85                  90

Ala Ile Trp Glu Lys Trp Gly Pro Pro Gln Glu Arg Ser Arg Leu
                 95                 100                 105

Cys Ser Ile Ala Leu Ser Gly Met Leu Leu Gly Cys Phe Thr Ala
                110                 115                 120

Ile Leu Ile Gly Gly Phe Ile Ser Glu Thr Leu Gly Trp Pro Phe
                125                 130                 135

Val Phe Tyr Ile Phe Gly Gly Val Gly Cys Val Cys Cys Leu Leu
                140                 145                 150

Trp Phe Val Val Ile Tyr Asp Asp Pro Val Ser Tyr Pro Trp Ile
                155                 160                 165

Ser Thr Ser Glu Lys Glu Tyr Ile Ile Ser Ser Leu Lys Gln Gln
                170                 175                 180

Val Arg Ser Ser Lys Gln Pro Leu Pro Ile Lys Ala Met Leu Arg
                185                 190                 195

Ser Leu Pro Ile Trp Ser Ile Cys Leu Gly Cys Phe Ser His Gln
                200                 205                 210

Trp Leu Val Ser Thr Met Val Val Tyr Ile Pro Thr Tyr Ile Ser
                215                 220                 225

Ser Val Tyr His Val Asn Ile Arg Asp Asn Gly Leu Leu Ser Ala
                230                 235                 240

Leu Pro Phe Ile Val Ala Trp Val Ile Gly Met Val Gly Gly Tyr
                245                 250                 255

Leu Ala Asp Phe Leu Leu Thr Lys Lys Phe Arg Leu Ile Thr Val
                260                 265                 270

Arg Lys Ile Ala Thr Ile Leu Gly Ser Leu Pro Ser Ser Ala Leu
                275                 280                 285

Ile Val Ser Leu Pro Tyr Leu Asn Ser Gly Tyr Ile Thr Ala Thr
                290                 295                 300

Ala Leu Leu Thr Leu Ser Cys Gly Leu Ser Thr Leu Cys Gln Ser
                305                 310                 315

Gly Ile Tyr Ile Asn Val Leu Asp Ile Ala Pro Arg Tyr Ser Ser
                320                 325                 330
```

```
 Phe Leu Met Gly Ala Ser Arg Gly Phe Ser Ser Ile Ala Pro Val
                 335                 340                 345

 Ile Val Pro Thr Val Ser Gly Phe Leu Leu Ser Gln Asp Pro Glu
                 350                 355                 360

 Phe Gly Trp Arg Asn Val Phe Phe Leu Leu Phe Ala Val Asn Leu
                 365                 370                 375

 Leu Gly Leu Leu Phe Tyr Leu Ile Phe Gly Glu Ala Asp Val Gln
                 380                 385                 390

 Glu Trp Ala Lys Glu Arg Lys Leu Thr Arg Leu
                 395                 400

<210> 114
<211> 507
<212> PRT
<213> Homo sapiens

<400> 114
 Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala
   1               5                  10                  15

 Ala Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys
                  20                  25                  30

 Ser Ala Asp Gly Ser Ala Pro Ala Gly Glu Gly Glu Gly Val Thr
                  35                  40                  45

 Leu Gln Arg Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val
                  50                  55                  60

 Gly Thr Ile Ile Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val
                  65                  70                  75

 Leu Lys Glu Ala Gly Ser Pro Gly Leu Ala Leu Val Val Trp Ala
                  80                  85                  90

 Ala Cys Gly Val Phe Ser Ile Val Gly Ala Leu Cys Tyr Ala Glu
                  95                 100                 105

 Leu Gly Thr Thr Ile Ser Lys Ser Gly Gly Asp Tyr Ala Tyr Met
                 110                 115                 120

 Leu Glu Val Tyr Gly Ser Leu Pro Ala Phe Leu Lys Leu Trp Ile
                 125                 130                 135

 Glu Leu Leu Ile Ile Arg Pro Ser Ser Gln Tyr Ile Val Ala Leu
                 140                 145                 150

 Val Phe Ala Thr Tyr Leu Leu Lys Pro Leu Phe Pro Thr Cys Pro
                 155                 160                 165

 Val Pro Glu Glu Ala Ala Lys Leu Val Ala Cys Leu Cys Val Leu
                 170                 175                 180

 Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys Ala Ala Thr Arg
                 185                 190                 195
```

```
Val Gln Asp Ala Phe Ala Ala Ala Lys Leu Leu Ala Leu Ala Leu
                200                 205                 210

Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp Val Ser
                215                 220                 225

Asn Leu Asp Pro Lys Phe Ser Phe Glu Gly Thr Lys Leu Asp Val
                230                 235                 240

Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly
                245                 250                 255

Gly Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro
                260                 265                 270

Tyr Arg Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val
                275                 280                 285

Thr Leu Val Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu
                290                 295                 300

Ser Thr Glu Gln Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe
                305                 310                 315

Gly Asn Tyr His Leu Gly Val Met Ser Trp Ile Ile Pro Val Phe
                320                 325                 330

Val Gly Leu Ser Cys Phe Gly Ser Val Asn Gly Ser Leu Phe Thr
                335                 340                 345

Ser Ser Arg Leu Phe Phe Val Gly Ser Arg Glu Gly His Leu Pro
                350                 355                 360

Ser Ile Leu Ser Met Ile His Pro Gln Leu Leu Thr Pro Val Pro
                365                 370                 375

Ser Leu Val Phe Thr Cys Val Met Thr Leu Leu Tyr Ala Phe Ser
                380                 385                 390

Lys Asp Ile Phe Ser Val Ile Asn Phe Phe Ser Phe Phe Asn Trp
                395                 400                 405

Leu Cys Val Ala Leu Ala Ile Ile Gly Met Ile Trp Leu Arg His
                410                 415                 420

Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val Asn Leu Ala Leu
                425                 430                 435

Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile Ala Val Ser
                440                 445                 450

Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr Ile Ile
                455                 460                 465

Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys Asn
                470                 475                 480

Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu
                485                 490                 495

Cys Gln Lys Leu Met Gln Val Val Pro Gln Glu Thr
```

500 505

<210> 115
<211> 507
<212> PRT
<213> Homo sapiens

<400> 115
Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala
1 5 10 15

Ala Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys
20 25 30

Ser Ala Asp Gly Ser Ala Pro Ala Gly Glu Gly Glu Gly Val Thr
35 40 45

Leu Gln Arg Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val
50 55 60

Gly Thr Ile Ile Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val
65 70 75

Leu Lys Glu Ala Gly Ser Pro Gly Leu Ala Leu Val Val Trp Ala
80 85 90

Ala Cys Gly Val Phe Ser Ile Val Gly Ala Leu Cys Tyr Ala Glu
95 100 105

Leu Gly Thr Thr Ile Ser Lys Ser Gly Gly Asp Tyr Ala Tyr Met
110 115 120

Leu Glu Val Tyr Gly Ser Leu Pro Ala Phe Leu Lys Leu Trp Ile
125 130 135

Glu Leu Leu Ile Ile Arg Pro Ser Ser Gln Tyr Ile Val Ala Leu
140 145 150

Val Phe Ala Thr Tyr Leu Leu Lys Pro Leu Phe Pro Thr Cys Pro
155 160 165

Val Pro Glu Glu Ala Ala Lys Leu Val Ala Cys Leu Cys Val Leu
170 175 180

Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys Ala Ala Thr Arg
185 190 195

Val Gln Asp Ala Phe Ala Ala Ala Lys Leu Leu Ala Leu Ala Leu
200 205 210

Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp Val Ser
215 220 225

Asn Leu Asp Pro Asn Phe Ser Phe Glu Gly Thr Lys Leu Asp Val
230 235 240

Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly
245 250 255

Gly Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro
260 265 270

```
 Tyr Arg Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val
                 275                 280                 285

 Thr Leu Val Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu
                 290                 295                 300

 Ser Thr Glu Gln Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe
                 305                 310                 315

 Gly Asn Tyr His Leu Gly Val Met Ser Trp Ile Ile Pro Val Phe
                 320                 325                 330

 Val Gly Leu Ser Cys Phe Gly Ser Val Asn Gly Ser Leu Phe Thr
                 335                 340                 345

 Ser Ser Arg Leu Phe Phe Val Gly Ser Arg Glu Gly His Leu Pro
                 350                 355                 360

 Ser Ile Leu Ser Met Ile His Pro Gln Leu Leu Thr Pro Val Pro
                 365                 370                 375

 Ser Leu Val Phe Thr Cys Val Met Thr Leu Leu Tyr Ala Phe Ser
                 380                 385                 390

 Lys Asp Ile Phe Ser Val Ile Asn Phe Phe Ser Phe Phe Asn Trp
                 395                 400                 405

 Leu Cys Val Ala Leu Ala Ile Ile Gly Met Ile Trp Leu Arg His
                 410                 415                 420

 Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val Asn Leu Ala Leu
                 425                 430                 435

 Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile Ala Val Ser
                 440                 445                 450

 Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr Ile Ile
                 455                 460                 465

 Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys Asn
                 470                 475                 480

 Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu
                 485                 490                 495

 Cys Gln Lys Leu Met Gln Val Val Pro Gln Glu Thr
                 500                 505

<210> 116
<211> 622
<212> PRT
<213> Homo sapiens

<400> 116
 Met Ala Leu Pro Thr Ala Arg Pro Leu Leu Gly Ser Cys Gly Thr
   1               5                  10                  15

 Pro Ala Leu Gly Ser Leu Leu Phe Leu Leu Phe Ser Leu Gly Trp
                  20                  25                  30
```

```
Val Gln Pro Ser Arg Thr Leu Ala Gly Glu Thr Gly Gln Glu Ala
                 35                  40                  45

Ala Pro Leu Asp Gly Val Leu Ala Asn Pro Pro Asn Ile Ser Ser
                 50                  55                  60

Leu Ser Pro Arg Gln Leu Leu Gly Phe Pro Cys Ala Glu Val Ser
                 65                  70                  75

Gly Leu Ser Thr Glu Arg Val Arg Glu Leu Ala Val Ala Leu Ala
                 80                  85                  90

Gln Lys Asn Val Lys Leu Ser Thr Glu Gln Leu Arg Cys Leu Ala
                 95                 100                 105

His Arg Leu Ser Glu Pro Pro Glu Asp Leu Asp Ala Leu Pro Leu
                110                 115                 120

Asp Leu Leu Leu Phe Leu Asn Pro Asp Ala Phe Ser Gly Pro Gln
                125                 130                 135

Ala Cys Thr Arg Phe Phe Ser Arg Ile Thr Lys Ala Asn Val Asp
                140                 145                 150

Leu Leu Pro Arg Gly Ala Pro Glu Arg Gln Arg Leu Leu Pro Ala
                155                 160                 165

Ala Leu Ala Cys Trp Gly Val Arg Gly Ser Leu Leu Ser Glu Ala
                170                 175                 180

Asp Val Arg Ala Leu Gly Gly Leu Ala Cys Asp Leu Pro Gly Arg
                185                 190                 195

Phe Val Ala Glu Ser Ala Glu Val Leu Leu Pro Arg Leu Val Ser
                200                 205                 210

Cys Pro Gly Pro Leu Asp Gln Asp Gln Gln Glu Ala Ala Arg Ala
                215                 220                 225

Ala Leu Gln Gly Gly Gly Pro Pro Tyr Gly Pro Pro Ser Thr Trp
                230                 235                 240

Ser Val Ser Thr Met Asp Ala Leu Arg Gly Leu Leu Pro Val Leu
                245                 250                 255

Gly Gln Pro Ile Ile Arg Ser Ile Pro Gln Gly Ile Val Ala Ala
                260                 265                 270

Trp Arg Gln Arg Ser Ser Arg Asp Pro Ser Trp Arg Gln Pro Glu
                275                 280                 285

Arg Thr Ile Leu Arg Pro Arg Phe Arg Arg Glu Val Glu Lys Thr
                290                 295                 300

Ala Cys Pro Ser Gly Lys Lys Ala Arg Glu Ile Asp Glu Ser Leu
                305                 310                 315

Ile Phe Tyr Lys Lys Trp Glu Leu Glu Ala Cys Val Asp Ala Ala
                320                 325                 330

Leu Leu Ala Thr Gln Met Asp Arg Val Asn Ala Ile Pro Phe Thr
```

```
                335                 340                 345

 Tyr Glu Gln Leu Asp Val Leu Lys His Lys Leu Asp Glu Leu Tyr
                350                 355                 360

 Pro Gln Gly Tyr Pro Glu Ser Val Ile Gln His Leu Gly Tyr Leu
                365                 370                 375

 Phe Leu Lys Met Ser Pro Glu Asp Ile Arg Lys Trp Asn Val Thr
                380                 385                 390

 Ser Leu Glu Thr Leu Lys Ala Leu Leu Glu Val Asn Lys Gly His
                395                 400                 405

 Glu Met Ser Pro Gln Val Ala Thr Leu Ile Asp Arg Phe Val Lys
                410                 415                 420

 Gly Arg Gly Gln Leu Asp Lys Asp Thr Leu Asp Thr Leu Thr Ala
                425                 430                 435

 Phe Tyr Pro Gly Tyr Leu Cys Ser Leu Ser Pro Glu Glu Leu Ser
                440                 445                 450

 Ser Val Pro Pro Ser Ser Ile Trp Ala Val Arg Pro Gln Asp Leu
                455                 460                 465

 Asp Thr Cys Asp Pro Arg Gln Leu Asp Val Leu Tyr Pro Lys Ala
                470                 475                 480

 Arg Leu Ala Phe Gln Asn Met Asn Gly Ser Glu Tyr Phe Val Lys
                485                 490                 495

 Ile Gln Ser Phe Leu Gly Gly Ala Pro Thr Glu Asp Leu Lys Ala
                500                 505                 510

 Leu Ser Gln Gln Asn Val Ser Met Asp Leu Ala Thr Phe Met Lys
                515                 520                 525

 Leu Arg Thr Asp Ala Val Leu Pro Leu Thr Val Ala Glu Val Gln
                530                 535                 540

 Lys Leu Leu Gly Pro His Val Glu Gly Leu Lys Ala Glu Glu Arg
                545                 550                 555

 His Arg Pro Val Arg Asp Trp Ile Leu Arg Gln Arg Gln Asp Asp
                560                 565                 570

 Leu Asp Thr Leu Gly Leu Gly Leu Gln Gly Gly Ile Pro Asn Gly
                575                 580                 585

 Tyr Leu Val Leu Asp Leu Ser Val Gln Glu Ala Leu Ser Gly Thr
                590                 595                 600

 Pro Cys Leu Leu Gly Pro Gly Pro Val Leu Thr Val Leu Ala Leu
                605                 610                 615

 Leu Leu Ala Ser Thr Leu Ala
                620

<210> 117
<211> 621
```

<212> PRT
<213> Homo sapiens

<400> 117
```
 Ala Leu Pro Thr Ala Arg Pro Leu Leu Gly Ser Cys Gly Thr Pro
   1               5                  10                  15

 Ala Leu Gly Ser Leu Leu Phe Leu Leu Phe Ser Leu Gly Trp Val
                20                  25                  30

 Gln Pro Ser Arg Thr Leu Ala Gly Glu Thr Gly Gln Glu Ala Ala
                35                  40                  45

 Pro Leu Asp Gly Val Leu Ala Asn Pro Pro Asn Ile Ser Ser Leu
                50                  55                  60

 Ser Pro Arg Gln Leu Leu Gly Phe Pro Cys Ala Glu Val Ser Gly
                65                  70                  75

 Leu Ser Thr Glu Arg Val Arg Glu Leu Ala Val Ala Leu Ala Gln
                80                  85                  90

 Lys Asn Val Lys Leu Ser Thr Glu Gln Leu Arg Cys Leu Ala His
                95                 100                 105

 Arg Leu Ser Glu Pro Pro Glu Asp Leu Asp Ala Leu Pro Leu Asp
               110                 115                 120

 Leu Leu Leu Phe Leu Asn Pro Asp Ala Phe Ser Gly Pro Gln Ala
               125                 130                 135

 Cys Thr Arg Phe Phe Ser Arg Ile Thr Lys Ala Asn Val Asp Leu
               140                 145                 150

 Leu Pro Arg Gly Ala Pro Glu Arg Gln Arg Leu Leu Pro Ala Ala
               155                 160                 165

 Leu Ala Cys Trp Gly Val Arg Gly Ser Leu Leu Ser Glu Ala Asp
               170                 175                 180

 Val Arg Ala Leu Gly Gly Leu Ala Cys Asp Leu Pro Gly Arg Phe
               185                 190                 195

 Val Ala Glu Ser Ala Glu Val Leu Leu Pro Arg Leu Val Ser Cys
               200                 205                 210

 Pro Gly Pro Leu Asp Gln Asp Gln Gln Glu Ala Ala Arg Ala Ala
               215                 220                 225

 Leu Gln Gly Gly Gly Pro Pro Tyr Gly Pro Pro Ser Thr Trp Ser
               230                 235                 240

 Val Ser Thr Met Asp Ala Leu Arg Gly Leu Leu Pro Val Leu Gly
               245                 250                 255

 Gln Pro Ile Ile Arg Ser Ile Pro Gln Gly Ile Val Ala Ala Trp
               260                 265                 270

 Arg Gln Arg Ser Ser Arg Asp Pro Ser Trp Arg Gln Pro Glu Arg
               275                 280                 285
```

```
Thr Ile Leu Arg Pro Arg Phe Arg Arg Glu Val Glu Lys Thr Ala
                290                 295                 300

Cys Pro Ser Gly Lys Lys Ala Arg Glu Ile Asp Glu Ser Leu Ile
                305                 310                 315

Phe Tyr Lys Lys Trp Glu Leu Glu Ala Cys Val Asp Ala Ala Leu
                320                 325                 330

Leu Ala Thr Gln Met Asp Arg Val Asn Ala Ile Pro Phe Thr Tyr
                335                 340                 345

Glu Gln Leu Asp Val Leu Lys His Lys Leu Asp Glu Leu Tyr Pro
                350                 355                 360

Gln Gly Tyr Pro Glu Ser Val Ile Gln His Leu Gly Tyr Leu Phe
                365                 370                 375

Leu Lys Met Ser Pro Glu Asp Ile Arg Lys Trp Asn Val Thr Ser
                380                 385                 390

Leu Glu Thr Leu Lys Ala Leu Leu Glu Val Asn Lys Gly His Glu
                395                 400                 405

Met Ser Pro Gln Val Ala Thr Leu Ile Asp Arg Phe Val Lys Gly
                410                 415                 420

Arg Gly Gln Leu Asp Lys Asp Thr Leu Asp Thr Leu Thr Ala Phe
                425                 430                 435

Tyr Pro Gly Tyr Leu Cys Ser Leu Ser Pro Glu Glu Leu Ser Ser
                440                 445                 450

Val Pro Pro Ser Ser Ile Trp Ala Val Arg Pro Gln Asp Leu Asp
                455                 460                 465

Thr Cys Asp Pro Arg Gln Leu Asp Val Leu Tyr Pro Lys Ala Arg
                470                 475                 480

Leu Ala Phe Gln Asn Met Asn Gly Ser Glu Tyr Phe Val Lys Ile
                485                 490                 495

Gln Ser Phe Leu Gly Gly Ala Pro Thr Glu Asp Leu Lys Ala Leu
                500                 505                 510

Ser Gln Gln Asn Val Ser Met Asp Leu Ala Thr Phe Met Lys Leu
                515                 520                 525

Arg Thr Asp Ala Val Leu Pro Leu Thr Val Ala Glu Val Gln Lys
                530                 535                 540

Leu Leu Gly Pro His Val Glu Gly Leu Lys Ala Glu Glu Arg His
                545                 550                 555

Arg Pro Val Arg Asp Trp Ile Leu Arg Gln Arg Gln Asp Asp Leu
                560                 565                 570

Asp Thr Leu Gly Leu Gly Leu Gln Gly Gly Ile Pro Asn Gly Tyr
                575                 580                 585

Leu Val Leu Asp Leu Ser Val Gln Glu Ala Leu Ser Gly Thr Pro
```

```
                590                 595                 600

 Cys Leu Leu Gly Pro Gly Pro Val Leu Thr Val Leu Ala Leu Leu
                605                 610                 615

 Leu Ala Ser Thr Leu Ala
                620
```

<210> 118
<211> 357
<212> PRT
<213> Homo sapiens

<400> 118

```
 Met Ala Thr Thr Val Pro Asp Gly Cys Arg Asn Gly Leu Lys Ser
   1               5                  10                  15

 Lys Tyr Tyr Arg Leu Cys Asp Lys Ala Glu Ala Trp Gly Ile Val
                20                  25                  30

 Leu Glu Thr Val Ala Thr Ala Gly Val Val Thr Ser Val Ala Phe
                35                  40                  45

 Met Leu Thr Leu Pro Ile Leu Val Cys Lys Val Gln Asp Ser Asn
                50                  55                  60

 Arg Arg Lys Met Leu Pro Thr Gln Phe Leu Phe Leu Leu Gly Val
                65                  70                  75

 Leu Gly Ile Phe Gly Leu Thr Phe Ala Phe Ile Ile Gly Leu Asp
                80                  85                  90

 Gly Ser Thr Gly Pro Thr Arg Phe Phe Leu Phe Gly Ile Leu Phe
                95                 100                 105

 Ser Ile Cys Phe Ser Cys Leu Leu Ala His Ala Val Ser Leu Thr
               110                 115                 120

 Lys Leu Val Arg Gly Arg Lys Pro Leu Ser Leu Leu Val Ile Leu
               125                 130                 135

 Gly Leu Ala Val Gly Phe Ser Leu Val Gln Asp Val Ile Ala Ile
               140                 145                 150

 Glu Tyr Ile Val Leu Thr Met Asn Arg Thr Asn Val Asn Val Phe
               155                 160                 165

 Ser Glu Leu Ser Ala Pro Arg Arg Asn Glu Asp Phe Val Leu Leu
               170                 175                 180

 Leu Thr Tyr Val Leu Phe Leu Met Ala Leu Thr Phe Leu Met Ser
               185                 190                 195

 Ser Phe Thr Phe Cys Gly Ser Phe Thr Gly Trp Lys Arg His Gly
               200                 205                 210

 Ala His Ile Tyr Leu Thr Met Leu Leu Ser Ile Ala Ile Trp Val
               215                 220                 225

 Ala Trp Ile Thr Leu Leu Met Leu Pro Asp Phe Asp Arg Arg Trp
               230                 235                 240
```

```
 Asp Asp Thr Ile Leu Ser Ser Ala Leu Ala Ala Asn Gly Trp Val
                 245                 250                 255

 Phe Leu Leu Ala Tyr Val Ser Pro Glu Phe Trp Leu Leu Thr Lys
                 260                 265                 270

 Gln Arg Asn Pro Met Asp Tyr Pro Val Glu Asp Ala Phe Cys Lys
                 275                 280                 285

 Pro Gln Leu Val Lys Lys Ser Tyr Gly Val Glu Asn Arg Ala Tyr
                 290                 295                 300

 Ser Gln Glu Glu Ile Thr Gln Gly Phe Glu Glu Thr Gly Asp Thr
                 305                 310                 315

 Leu Tyr Ala Pro Tyr Ser Thr His Phe Gln Leu Gln Asn Gln Pro
                 320                 325                 330

 Pro Gln Lys Glu Phe Ser Ile Pro Arg Ala His Ala Trp Pro Ser
                 335                 340                 345

 Pro Tyr Lys Asp Tyr Glu Val Lys Lys Glu Gly Ser
                 350                 355

<210> 119
<211> 836
<212> PRT
<213> Homo sapiens

<400> 119
 Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
   1               5                  10                  15

 Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
                  20                  25                  30

 His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
                  35                  40                  45

 Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
                  50                  55                  60

 Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
                  65                  70                  75

 Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
                  80                  85                  90

 Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
                  95                 100                 105

 Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
                 110                 115                 120

 Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
                 125                 130                 135

 Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
                 140                 145                 150
```

```
Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
                155                 160                 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
                170                 175                 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
                185                 190                 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
                200                 205                 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
                215                 220                 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
                230                 235                 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
                245                 250                 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
                260                 265                 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
                275                 280                 285

Val Leu Glu Arg Phe Met Gly Asp Lys Val Ala Ser Glu Ala Leu
                290                 295                 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
                305                 310                 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
                320                 325                 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
                335                 340                 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
                350                 355                 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
                365                 370                 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
                380                 385                 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
                395                 400                 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
                410                 415                 420

Val Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
                425                 430                 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
                440                 445                 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
```

```
                455                 460                 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
                470                 475                 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
                485                 490                 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
                500                 505                 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
                515                 520                 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
                530                 535                 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
                545                 550                 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
                560                 565                 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
                575                 580                 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
                590                 595                 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
                605                 610                 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
                620                 625                 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
                635                 640                 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
                650                 655                 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Thr Thr Lys Ile Ile Thr
                665                 670                 675

Lys Val Val Glu Pro Lys Ile Lys Val Ile Glu Gly Ser Leu Gln
                680                 685                 690

Pro Ile Ile Lys Thr Glu Gly Pro Thr Leu Thr Lys Val Lys Ile
                695                 700                 705

Glu Gly Glu Pro Glu Phe Arg Leu Ile Lys Glu Gly Glu Thr Ile
                710                 715                 720

Thr Glu Val Ile His Gly Glu Pro Ile Ile Lys Lys Tyr Thr Lys
                725                 730                 735

Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys Glu Thr Arg
                740                 745                 750

Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr Thr Arg Ile
                755                 760                 765
```

```
 Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys Lys Leu Leu
                 770                 775                 780

 Gln Glu Glu Val Thr Lys Val Thr Lys Phe Ile Glu Gly Gly Asp
                 785                 790                 795

 Gly His Leu Phe Glu Asp Glu Glu Ile Lys Arg Leu Leu Gln Gly
                 800                 805                 810

 Asp Thr Pro Val Arg Lys Leu Gln Ala Asn Lys Lys Val Gln Gly
                 815                 820                 825

 Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser Gln
                 830                 835

<210> 120
<211> 809
<212> PRT
<213> Homo sapiens

<400> 120
 Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
   1               5                  10                  15

 Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
                  20                  25                  30

 His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
                  35                  40                  45

 Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
                  50                  55                  60

 Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
                  65                  70                  75

 Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
                  80                  85                  90

 Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
                  95                 100                 105

 Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
                 110                 115                 120

 Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
                 125                 130                 135

 Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
                 140                 145                 150

 Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
                 155                 160                 165

 His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
                 170                 175                 180

 Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
                 185                 190                 195
```

```
Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
                200                 205                 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
                215                 220                 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
                230                 235                 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
                245                 250                 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
                260                 265                 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
                275                 280                 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
                290                 295                 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
                305                 310                 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
                320                 325                 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
                335                 340                 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
                350                 355                 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
                365                 370                 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
                380                 385                 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
                395                 400                 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
                410                 415                 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
                425                 430                 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
                440                 445                 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
                455                 460                 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
                470                 475                 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
                485                 490                 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
```

```
                500                 505                 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
                515                 520                 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
                530                 535                 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
                545                 550                 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
                560                 565                 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
                575                 580                 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
                590                 595                 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
                605                 610                 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
                620                 625                 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
                635                 640                 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
                650                 655                 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Arg Pro Thr Leu Thr Lys
                665                 670                 675

Val Lys Ile Glu Gly Glu Pro Glu Phe Arg Leu Ile Lys Glu Gly
                680                 685                 690

Glu Thr Ile Thr Glu Val Ile His Gly Glu Pro Ile Ile Lys Lys
                695                 700                 705

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
                710                 715                 720

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
                725                 730                 735

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
                740                 745                 750

Lys Leu Leu Gln Glu Glu Val Thr Lys Val Thr Lys Phe Ile Glu
                755                 760                 765

Gly Gly Asp Gly His Leu Phe Glu Asp Glu Glu Ile Lys Arg Leu
                770                 775                 780

Leu Gln Gly Asp Thr Pro Val Arg Lys Leu Gln Ala Asn Lys Lys
                785                 790                 795

Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser Gln
                800                 805
```

<210> 121
<211> 751
<212> PRT
<213> Homo sapiens

<400> 121
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

```
Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
                275                 280                 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
                290                 295                 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
                305                 310                 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
                320                 325                 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
                335                 340                 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
                350                 355                 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
                365                 370                 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
                380                 385                 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
                395                 400                 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
                410                 415                 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
                425                 430                 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
                440                 445                 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
                455                 460                 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
                470                 475                 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
                485                 490                 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
                500                 505                 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
                515                 520                 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
                530                 535                 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
                545                 550                 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
                560                 565                 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
```

```
                575                 580                 585

 Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
                590                 595                 600

 Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
                605                 610                 615

 Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
                620                 625                 630

 Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
                635                 640                 645

 Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
                650                 655                 660

 Phe Lys Glu Ile Pro Val Thr Val Tyr Lys Pro Ile Ile Lys Lys
                665                 670                 675

 Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
                680                 685                 690

 Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
                695                 700                 705

 Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
                710                 715                 720

 Lys Leu Leu Gln Glu Asp Thr Pro Val Arg Lys Leu Gln Ala Asn
                725                 730                 735

 Lys Lys Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser
                740                 745                 750

 Gln


<210> 122
<211> 781
<212> PRT
<213> Homo sapiens

<400> 122
 Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
   1               5                  10                  15

 Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
                 20                  25                  30

 His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
                 35                  40                  45

 Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
                 50                  55                  60

 Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
                 65                  70                  75

 Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
                 80                  85                  90
```

```
Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
                 95                  100                 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
                110                 115                 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
                125                 130                 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
                140                 145                 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
                155                 160                 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
                170                 175                 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
                185                 190                 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
                200                 205                 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
                215                 220                 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
                230                 235                 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
                245                 250                 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
                260                 265                 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
                275                 280                 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
                290                 295                 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
                305                 310                 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
                320                 325                 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
                335                 340                 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
                350                 355                 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
                365                 370                 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
                380                 385                 390
```

```
Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
                395                 400                 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
                410                 415                 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
                425                 430                 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
                440                 445                 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
                455                 460                 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
                470                 475                 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
                485                 490                 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
                500                 505                 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
                515                 520                 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
                530                 535                 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
                545                 550                 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
                560                 565                 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
                575                 580                 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
                590                 595                 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
                605                 610                 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
                620                 625                 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
                635                 640                 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
                650                 655                 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Arg Pro Thr Leu Thr Lys
                665                 670                 675

Val Lys Ile Glu Gly Glu Pro Glu Phe Arg Leu Ile Lys Glu Gly
                680                 685                 690

Glu Thr Ile Thr Glu Val Ile His Gly Glu Pro Ile Ile Lys Lys
```

695 700 705

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
710 715 720

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
725 730 735

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
740 745 750

Lys Leu Leu Gln Glu Asp Thr Pro Val Arg Lys Leu Gln Ala Asn
755 760 765

Lys Lys Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser
770 775 780

Gln

<210> 123
<211> 779
<212> PRT
<213> Homo sapiens

<400> 123
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

```
Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
                185                 190                 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
                200                 205                 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
                215                 220                 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
                230                 235                 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
                245                 250                 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
                260                 265                 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
                275                 280                 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
                290                 295                 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
                305                 310                 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
                320                 325                 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
                335                 340                 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
                350                 355                 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
                365                 370                 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
                380                 385                 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
                395                 400                 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
                410                 415                 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
                425                 430                 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
                440                 445                 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
                455                 460                 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
                470                 475                 480
```

```
Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
                485                 490                 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
                500                 505                 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
                515                 520                 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
                530                 535                 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
                545                 550                 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
                560                 565                 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
                575                 580                 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
                590                 595                 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
                605                 610                 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
                620                 625                 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
                635                 640                 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
                650                 655                 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Lys Pro Ile Ile Lys Lys
                665                 670                 675

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
                680                 685                 690

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
                695                 700                 705

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
                710                 715                 720

Lys Leu Leu Gln Glu Glu Val Thr Lys Val Thr Lys Phe Ile Glu
                725                 730                 735

Gly Gly Asp Gly His Leu Phe Glu Asp Glu Glu Ile Lys Arg Leu
                740                 745                 750

Leu Gln Gly Asp Thr Pro Val Arg Lys Leu Gln Ala Asn Lys Lys
                755                 760                 765

Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser Gln
                770                 775
```

<210> 124

<211> 283
<212> PRT
<213> Homo sapiens

<400> 124
```
 Met Val Asn Tyr Ala Trp Ala Gly Arg Ser Gln Arg Lys Leu Trp
   1               5                  10                  15

 Trp Arg Ser Val Ala Val Leu Thr Cys Lys Ser Val Val Arg Pro
                  20                  25                  30

 Gly Tyr Arg Gly Gly Leu Gln Ala Arg Arg Ser Thr Leu Leu Lys
                  35                  40                  45

 Thr Cys Ala Arg Ala Arg Ala Thr Ala Pro Gly Ala Met Lys Met
                  50                  55                  60

 Val Ala Pro Trp Thr Arg Phe Tyr Ser Asn Ser Cys Cys Leu Cys
                  65                  70                  75

 Cys His Val Arg Thr Gly Thr Ile Leu Leu Gly Val Trp Tyr Leu
                  80                  85                  90

 Ile Ile Asn Ala Val Val Leu Leu Ile Leu Leu Ser Ala Leu Ala
                  95                 100                 105

 Asp Pro Asp Gln Tyr Asn Phe Ser Ser Ser Glu Leu Gly Gly Asp
                 110                 115                 120

 Phe Glu Phe Met Asp Asp Ala Asn Met Cys Ile Ala Ile Ala Ile
                 125                 130                 135

 Ser Leu Leu Met Ile Leu Ile Cys Ala Met Ala Thr Tyr Gly Ala
                 140                 145                 150

 Tyr Lys Gln Arg Ala Ala Trp Ile Ile Pro Phe Phe Cys Tyr Gln
                 155                 160                 165

 Ile Phe Asp Phe Ala Leu Asn Met Leu Val Ala Ile Thr Val Leu
                 170                 175                 180

 Ile Tyr Pro Asn Ser Ile Gln Glu Tyr Ile Arg Gln Leu Pro Pro
                 185                 190                 195

 Asn Phe Pro Tyr Arg Asp Asp Val Met Ser Val Asn Pro Thr Cys
                 200                 205                 210

 Leu Val Leu Ile Ile Leu Leu Phe Ile Ser Ile Ile Leu Thr Phe
                 215                 220                 225

 Lys Gly Tyr Leu Ile Ser Cys Val Trp Asn Cys Tyr Arg Tyr Ile
                 230                 235                 240

 Asn Gly Arg Asn Ser Ser Asp Val Leu Val Tyr Val Thr Ser Asn
                 245                 250                 255

 Asp Thr Thr Val Leu Leu Pro Pro Tyr Asp Asp Ala Thr Val Asn
                 260                 265                 270

 Gly Ala Ala Lys Glu Pro Pro Pro Pro Tyr Val Ser Ala
                 275                 280
```

<210> 125
<211> 254
<212> PRT
<213> Homo sapiens

<400> 125
Met Pro Gly Gln Gly Gly Leu Gln Ala Arg Arg Ser Thr Leu Leu
1 5 10 15

Lys Thr Cys Ala Arg Ala Arg Ala Thr Ala Pro Gly Ala Met Lys
20 25 30

Met Val Ala Pro Trp Thr Arg Phe Tyr Ser Asn Ser Cys Cys Leu
35 40 45

Cys Cys His Val Arg Thr Gly Thr Ile Leu Leu Gly Val Trp Tyr
50 55 60

Leu Ile Ile Asn Ala Val Val Leu Leu Ile Leu Leu Ser Ala Leu
65 70 75

Ala Asp Pro Asp Gln Tyr Asn Phe Ser Ser Ser Glu Leu Gly Gly
80 85 90

Asp Phe Glu Phe Met Asp Asp Ala Asn Met Cys Ile Ala Ile Ala
95 100 105

Ile Ser Leu Leu Met Ile Leu Ile Cys Ala Met Ala Thr Tyr Gly
110 115 120

Ala Tyr Lys Gln Arg Ala Ala Trp Ile Ile Pro Phe Phe Cys Tyr
125 130 135

Gln Ile Phe Asp Phe Ala Leu Asn Met Leu Val Ala Ile Thr Val
140 145 150

Leu Ile Tyr Pro Asn Ser Ile Gln Glu Tyr Ile Arg Gln Leu Pro
155 160 165

Pro Asn Phe Pro Tyr Arg Asp Asp Val Met Ser Val Asn Pro Thr
170 175 180

Cys Leu Val Leu Ile Ile Leu Leu Phe Ile Ser Ile Ile Leu Thr
185 190 195

Phe Lys Gly Tyr Leu Ile Ser Cys Val Trp Asn Cys Tyr Arg Tyr
200 205 210

Ile Asn Gly Arg Asn Ser Ser Asp Val Leu Val Tyr Val Thr Ser
215 220 225

Asn Asp Thr Thr Val Leu Leu Pro Pro Tyr Asp Asp Ala Thr Val
230 235 240

Asn Gly Ala Ala Lys Glu Pro Pro Pro Pro Tyr Val Ser Ala
245 250

<210> 126
<211> 539
<212> PRT

<213> Homo sapiens

<400> 126

```
 Met Gly Arg Ala Gly Gly Gly Gly Pro Gly Arg Gly Pro Pro Pro
   1               5                  10                  15

 Leu Leu Leu Phe Leu Gly Ala Ala Leu Val Leu Ala Ser Gly Ala
                20                  25                  30

 Val Pro Ala Arg Glu Ala Gly Ser Ala Val Glu Ala Glu Glu Leu
                35                  40                  45

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

 Glu Glu Ala Gly Pro Pro Ala Ala Gly Glu Asp Glu Ala Ser Trp
                65                  70                  75

 Thr Ala Pro Gly Gly Glu Leu Ala Gly Pro Glu Glu Val Leu Gln
                80                  85                  90

 Glu Ser Ala Ala Val Thr Gly Thr Ala Trp Leu Glu Ala Asp Ser
                95                 100                 105

 Pro Gly Leu Gly Gly Val Thr Ala Glu Ala Gly Ser Gly Asp Ala
               110                 115                 120

 Gln Ala Leu Pro Ala Thr Leu Gln Ala Pro His Glu Val Leu Gly
               125                 130                 135

 Gln Ser Ile Met Pro Pro Ala Ile Pro Glu Ala Thr Glu Ala Ser
               140                 145                 150

 Gly Pro Pro Ser Pro Thr Pro Gly Asp Lys Leu Ser Pro Ala Ser
               155                 160                 165

 Glu Leu Pro Lys Glu Ser Pro Leu Glu Val Trp Leu Asn Leu Gly
               170                 175                 180

 Gly Ser Thr Pro Asp Pro Gln Val Pro Glu Leu Thr Tyr Pro Phe
               185                 190                 195

 Gln Gly Thr Leu Glu Pro Gln Pro Ala Ser Asp Ile Ile Asp Ile
               200                 205                 210

 Asp Tyr Phe Glu Gly Leu Asp Gly Glu Gly Arg Gly Ala Asp Leu
               215                 220                 225

 Gly Ser Phe Pro Gly Ser Pro Gly Thr Ser Glu Asn His Pro Asp
               230                 235                 240

 Thr Glu Gly Glu Thr Pro Ser Trp Ser Leu Leu Asp Leu Tyr Asp
               245                 250                 255

 Asp Phe Thr Pro Phe Asp Glu Ser Asp Phe Tyr Pro Thr Thr Ser
               260                 265                 270

 Phe Tyr Asp Asp Leu Asp Glu Glu Glu Glu Glu Glu Glu Asp Asp
               275                 280                 285

 Lys Asp Ala Val Gly Gly Gly Asp Leu Glu Asp Glu Asn Glu Leu
```

```
                290                 295                 300

 Leu Val Pro Thr Gly Lys Pro Gly Leu Gly Pro Gly Thr Gly Gln
                305                 310                 315

 Pro Thr Ser Arg Trp His Ala Val Pro Pro Gln His Thr Leu Gly
                320                 325                 330

 Ser Val Pro Gly Ser Ser Ile Ala Leu Arg Pro Arg Pro Gly Glu
                335                 340                 345

 Pro Gly Arg Asp Leu Ala Ser Ser Glu Asn Gly Thr Glu Cys Arg
                350                 355                 360

 Ser Gly Phe Val Arg His Asn Gly Ser Cys Arg Ser Val Cys Asp
                365                 370                 375

 Leu Phe Pro Ser Tyr Cys His Asn Gly Gly Gln Cys Tyr Leu Val
                380                 385                 390

 Glu Asn Ile Gly Ala Phe Cys Arg Cys Asn Thr Gln Asp Tyr Ile
                395                 400                 405

 Trp His Lys Gly Met Arg Cys Glu Ser Ile Ile Thr Asp Phe Gln
                410                 415                 420

 Val Met Cys Val Ala Val Gly Ser Ala Ala Leu Val Leu Leu Leu
                425                 430                 435

 Leu Phe Met Met Thr Val Phe Phe Ala Lys Lys Leu Tyr Leu Leu
                440                 445                 450

 Lys Thr Glu Asn Thr Lys Leu Arg Arg Thr Asn Lys Phe Arg Thr
                455                 460                 465

 Pro Ser Glu Leu His Asn Asp Asn Phe Ser Leu Ser Thr Ile Ala
                470                 475                 480

 Glu Gly Ser His Pro Asn Asp Asp Pro Ser Ala Pro His Lys Ile
                485                 490                 495

 Gln Glu Val Leu Lys Ser Cys Leu Lys Glu Glu Glu Ser Phe Asn
                500                 505                 510

 Ile Gln Asn Ser Met Ser Pro Lys Leu Glu Gly Gly Lys Gly Asp
                515                 520                 525

 Gln Ala Asp Leu Asp Val Asn Cys Leu Gln Asn Asn Leu Thr
                530                 535

<210> 127
<211> 1325
<212> PRT
<213> Homo sapiens

<400> 127
 Met Leu Pro Val Tyr Gln Glu Val Lys Pro Asn Pro Leu Gln Asp
   1               5                  10                  15

 Ala Asn Ile Cys Ser Arg Val Phe Phe Trp Trp Leu Asn Pro Leu
                 20                  25                  30
```

```
Phe Lys Ile Gly His Lys Arg Arg Leu Glu Glu Asp Asp Met Tyr
                 35                  40                  45

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

Gln Gly Phe Trp Asp Lys Glu Val Leu Arg Ala Glu Asn Asp Ala
                 65                  70                  75

Gln Lys Pro Ser Leu Thr Arg Ala Ile Ile Lys Cys Tyr Trp Lys
                 80                  85                  90

Ser Tyr Leu Val Leu Gly Ile Phe Thr Leu Ile Glu Glu Ser Ala
                 95                 100                 105

Lys Val Ile Gln Pro Ile Phe Leu Gly Lys Ile Ile Asn Tyr Phe
                110                 115                 120

Glu Asn Tyr Asp Pro Met Asp Ser Val Ala Leu Asn Thr Ala Tyr
                125                 130                 135

Ala Tyr Ala Thr Val Leu Thr Phe Cys Thr Leu Ile Leu Ala Ile
                140                 145                 150

Leu His His Leu Tyr Phe Tyr His Val Gln Cys Ala Gly Met Arg
                155                 160                 165

Leu Arg Val Ala Met Cys His Met Ile Tyr Arg Lys Ala Leu Arg
                170                 175                 180

Leu Ser Asn Met Ala Met Gly Lys Thr Thr Thr Gly Gln Ile Val
                185                 190                 195

Asn Leu Leu Ser Asn Asp Val Asn Lys Phe Asp Gln Val Thr Val
                200                 205                 210

Phe Leu His Phe Leu Trp Ala Gly Pro Leu Gln Ala Ile Ala Val
                215                 220                 225

Thr Ala Leu Leu Trp Met Glu Ile Gly Ile Ser Cys Leu Ala Gly
                230                 235                 240

Met Ala Val Leu Ile Ile Leu Leu Pro Leu Gln Ser Cys Phe Gly
                245                 250                 255

Lys Leu Phe Ser Ser Leu Arg Ser Lys Thr Ala Thr Phe Thr Asp
                260                 265                 270

Ala Arg Ile Arg Thr Met Asn Glu Val Ile Thr Gly Ile Arg Ile
                275                 280                 285

Ile Lys Met Tyr Ala Trp Glu Lys Ser Phe Ser Asn Leu Ile Thr
                290                 295                 300

Asn Leu Arg Lys Lys Glu Ile Ser Lys Ile Leu Arg Ser Ser Cys
                305                 310                 315

Leu Arg Gly Met Asn Leu Ala Ser Phe Phe Ser Ala Ser Lys Ile
                320                 325                 330
```

```
Ile Val Phe Val Thr Phe Thr Thr Tyr Val Leu Leu Gly Ser Val
                335                 340                 345

Ile Thr Ala Ser Arg Val Phe Val Ala Val Thr Leu Tyr Gly Ala
                350                 355                 360

Val Arg Leu Thr Val Thr Leu Phe Phe Pro Ser Ala Ile Glu Arg
                365                 370                 375

Val Ser Glu Ala Ile Val Ser Ile Arg Arg Ile Gln Thr Phe Leu
                380                 385                 390

Leu Leu Asp Glu Ile Ser Gln Arg Asn Arg Gln Leu Pro Ser Asp
                395                 400                 405

Gly Lys Lys Met Val His Val Gln Asp Phe Thr Ala Phe Trp Asp
                410                 415                 420

Lys Ala Ser Glu Thr Pro Thr Leu Gln Gly Leu Ser Phe Thr Val
                425                 430                 435

Arg Pro Gly Glu Leu Leu Ala Val Val Gly Pro Val Gly Ala Gly
                440                 445                 450

Lys Ser Ser Leu Leu Ser Ala Val Leu Gly Glu Leu Ala Pro Ser
                455                 460                 465

His Gly Leu Val Ser Val His Gly Arg Ile Ala Tyr Val Ser Gln
                470                 475                 480

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

Gly Lys Lys Tyr Glu Lys Glu Arg Tyr Glu Lys Val Ile Lys Ala
                500                 505                 510

Cys Ala Leu Lys Lys Asp Leu Gln Leu Leu Glu Asp Gly Asp Leu
                515                 520                 525

Thr Val Ile Gly Asp Arg Gly Thr Thr Leu Ser Gly Gly Gln Lys
                530                 535                 540

Ala Arg Val Asn Leu Ala Arg Ala Val Tyr Gln Asp Ala Asp Ile
                545                 550                 555

Tyr Leu Leu Asp Asp Pro Leu Ser Ala Val Asp Ala Glu Val Ser
                560                 565                 570

Arg His Leu Phe Glu Leu Cys Ile Cys Gln Ile Leu His Glu Lys
                575                 580                 585

Ile Thr Ile Leu Val Thr His Gln Leu Gln Tyr Leu Lys Ala Ala
                590                 595                 600

Ser Gln Ile Leu Ile Leu Lys Asp Gly Lys Met Val Gln Lys Gly
                605                 610                 615

Thr Tyr Thr Glu Phe Leu Lys Ser Gly Ile Asp Phe Gly Ser Leu
                620                 625                 630

Leu Lys Lys Asp Asn Glu Glu Ser Glu Gln Pro Pro Val Pro Gly
```

635 640 645

Thr Pro Thr Leu Arg Asn Arg Thr Phe Ser Glu Ser Ser Val Trp
650 655 660

Ser Gln Gln Ser Ser Arg Pro Ser Leu Lys Asp Gly Ala Leu Glu
665 670 675

Ser Gln Asp Thr Glu Asn Val Pro Val Thr Leu Ser Glu Glu Asn
680 685 690

Arg Ser Glu Gly Lys Val Gly Phe Gln Ala Tyr Lys Asn Tyr Phe
695 700 705

Arg Ala Gly Ala His Trp Ile Val Phe Ile Phe Leu Ile Leu Leu
710 715 720

Asn Thr Ala Ala Gln Val Ala Tyr Val Leu Gln Asp Trp Trp Leu
725 730 735

Ser Tyr Trp Ala Asn Lys Gln Ser Met Leu Asn Val Thr Val Asn
740 745 750

Gly Gly Gly Asn Val Thr Glu Lys Leu Asp Leu Asn Trp Tyr Leu
755 760 765

Gly Ile Tyr Ser Gly Leu Thr Val Ala Thr Val Leu Phe Gly Ile
770 775 780

Ala Arg Ser Leu Leu Val Phe Tyr Val Leu Val Asn Ser Ser Gln
785 790 795

Thr Leu His Asn Lys Met Phe Glu Ser Ile Leu Lys Ala Pro Val
800 805 810

Leu Phe Phe Asp Arg Asn Pro Ile Gly Arg Ile Leu Asn Arg Phe
815 820 825

Ser Lys Asp Ile Gly His Leu Asp Asp Leu Leu Pro Leu Thr Phe
830 835 840

Leu Asp Phe Ile Gln Thr Leu Leu Gln Val Val Gly Val Val Ser
845 850 855

Val Ala Val Ala Val Ile Pro Trp Ile Ala Ile Pro Leu Val Pro
860 865 870

Leu Gly Ile Ile Phe Ile Phe Leu Arg Arg Tyr Phe Leu Glu Thr
875 880 885

Ser Arg Asp Val Lys Arg Leu Glu Ser Thr Thr Arg Ser Pro Val
890 895 900

Phe Ser His Leu Ser Ser Ser Leu Gln Gly Leu Trp Thr Ile Arg
905 910 915

Ala Tyr Lys Ala Glu Glu Arg Cys Gln Glu Leu Phe Asp Ala His
920 925 930

Gln Asp Leu His Ser Glu Ala Trp Phe Leu Phe Leu Thr Thr Ser
935 940 945

```
Arg Trp Phe Ala Val Arg Leu Asp Ala Ile Cys Ala Met Phe Val
                950                 955                 960

Ile Ile Val Ala Phe Gly Ser Leu Ile Leu Ala Lys Thr Leu Asp
                965                 970                 975

Ala Gly Gln Val Gly Leu Ala Leu Ser Tyr Ala Leu Thr Leu Met
                980                 985                 990

Gly Met Phe Gln Trp Cys Val Arg Gln Ser Ala Glu Val Glu Asn
                995                1000                1005

Met Met Ile Ser Val Glu Arg Val Ile Glu Tyr Thr Asp Leu Glu
               1010                1015                1020

Lys Glu Ala Pro Trp Glu Tyr Gln Lys Arg Pro Pro Pro Ala Trp
               1025                1030                1035

Pro His Glu Gly Val Ile Ile Phe Asp Asn Val Asn Phe Met Tyr
               1040                1045                1050

Ser Pro Gly Gly Pro Leu Val Leu Lys His Leu Thr Ala Leu Ile
               1055                1060                1065

Lys Ser Gln Glu Lys Val Gly Ile Val Gly Arg Thr Gly Ala Gly
               1070                1075                1080

Lys Ser Ser Leu Ile Ser Ala Leu Phe Arg Leu Ser Glu Pro Glu
               1085                1090                1095

Gly Lys Ile Trp Ile Asp Lys Ile Leu Thr Thr Glu Ile Gly Leu
               1100                1105                1110

His Asp Leu Arg Lys Lys Met Ser Ile Ile Pro Gln Glu Pro Val
               1115                1120                1125

Leu Phe Thr Gly Thr Met Arg Lys Asn Leu Asp Pro Phe Lys Glu
               1130                1135                1140

His Thr Asp Glu Glu Leu Trp Asn Ala Leu Gln Glu Val Gln Leu
               1145                1150                1155

Lys Glu Thr Ile Glu Asp Leu Pro Gly Lys Met Asp Thr Glu Leu
               1160                1165                1170

Ala Glu Ser Gly Ser Asn Phe Ser Val Gly Gln Arg Gln Leu Val
               1175                1180                1185

Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn Gln Ile Leu Ile Ile
               1190                1195                1200

Asp Glu Ala Thr Ala Asn Val Asp Pro Arg Thr Asp Glu Leu Ile
               1205                1210                1215

Gln Lys Lys Ile Arg Glu Lys Phe Ala His Cys Thr Val Leu Thr
               1220                1225                1230

Ile Ala His Arg Leu Asn Thr Ile Ile Asp Ser Asp Lys Ile Met
               1235                1240                1245
```

```
 Val Leu Asp Ser Gly Arg Leu Lys Glu Tyr Asp Glu Pro Tyr Val
                1250                1255                1260

 Leu Leu Gln Asn Lys Glu Ser Leu Phe Tyr Lys Met Val Gln Gln
                1265                1270                1275

 Leu Gly Lys Ala Glu Ala Ala Ala Leu Thr Glu Thr Ala Lys Gln
                1280                1285                1290

 Val Tyr Phe Lys Arg Asn Tyr Pro His Ile Gly His Thr Asp His
                1295                1300                1305

 Met Val Thr Asn Thr Ser Asn Gly Gln Pro Ser Thr Leu Thr Ile
                1310                1315                1320

 Phe Glu Thr Ala Leu
                1325

<210> 128
<211> 501
<212> PRT
<213> Homo sapiens

<400> 128
 Met Ser Ala Leu Arg Arg Lys Phe Gly Asp Asp Tyr Gln Val Val
   1               5                  10                  15

 Thr Thr Ser Ser Ser Gly Ser Gly Leu Gln Pro Gln Gly Pro Gly
                20                  25                  30

 Gln Asp Pro Gln Gln Gln Leu Val Pro Lys Lys Lys Arg Gln Arg
                35                  40                  45

 Phe Val Asp Lys Asn Gly Arg Cys Asn Val Gln His Gly Asn Leu
                50                  55                  60

 Gly Ser Glu Thr Ser Arg Tyr Leu Ser Asp Leu Phe Thr Thr Leu
                65                  70                  75

 Val Asp Leu Lys Trp Arg Trp Asn Leu Phe Ile Phe Ile Leu Thr
                80                  85                  90

 Tyr Thr Val Ala Trp Leu Phe Met Ala Ser Met Trp Trp Val Ile
                95                 100                 105

 Ala Tyr Thr Arg Gly Asp Leu Asn Lys Ala His Val Gly Asn Tyr
                110                 115                 120

 Thr Pro Cys Val Ala Asn Val Tyr Asn Phe Pro Ser Ala Phe Leu
                125                 130                 135

 Phe Phe Ile Glu Thr Glu Ala Thr Ile Gly Tyr Gly Tyr Arg Tyr
                140                 145                 150

 Ile Thr Asp Lys Cys Pro Glu Gly Ile Ile Leu Phe Leu Phe Gln
                155                 160                 165

 Ser Ile Leu Gly Ser Ile Val Asp Ala Phe Leu Ile Gly Cys Met
                170                 175                 180

 Phe Ile Lys Met Ser Gln Pro Lys Lys Arg Ala Glu Thr Leu Met
```

```
                185                 190                 195

Phe Ser Glu His Ala Val Ile Ser Met Arg Asp Gly Lys Leu Thr
                200                 205                 210

Leu Met Phe Arg Val Gly Asn Leu Arg Asn Ser His Met Val Ser
                215                 220                 225

Ala Gln Ile Arg Cys Lys Leu Leu Lys Ser Arg Gln Thr Pro Glu
                230                 235                 240

Gly Glu Phe Leu Pro Leu Asp Gln Leu Glu Leu Asp Val Gly Phe
                245                 250                 255

Ser Thr Gly Ala Asp Gln Leu Phe Leu Val Ser Pro Leu Thr Ile
                260                 265                 270

Cys His Val Ile Asp Ala Lys Ser Pro Phe Tyr Asp Leu Ser Gln
                275                 280                 285

Arg Ser Met Gln Thr Glu Gln Phe Glu Ile Val Val Ile Leu Glu
                290                 295                 300

Gly Ile Val Glu Thr Thr Gly Met Thr Cys Gln Ala Arg Thr Ser
                305                 310                 315

Tyr Thr Glu Asp Glu Val Leu Trp Gly His Arg Phe Phe Pro Val
                320                 325                 330

Ile Ser Leu Glu Glu Gly Phe Phe Lys Val Asp Tyr Ser Gln Phe
                335                 340                 345

His Ala Thr Phe Glu Val Pro Thr Pro Pro Tyr Ser Val Lys Glu
                350                 355                 360

Gln Glu Glu Met Leu Leu Met Ser Ser Pro Leu Ile Ala Pro Ala
                365                 370                 375

Ile Thr Asn Ser Lys Glu Arg His Asn Ser Val Glu Cys Leu Asp
                380                 385                 390

Gly Leu Asp Asp Ile Thr Thr Lys Leu Pro Ser Lys Leu Gln Lys
                395                 400                 405

Ile Thr Gly Arg Glu Asp Phe Pro Lys Lys Leu Leu Arg Met Ser
                410                 415                 420

Ser Thr Thr Ser Glu Lys Ala Tyr Ser Leu Gly Asp Leu Pro Met
                425                 430                 435

Lys Leu Gln Arg Ile Ser Ser Val Pro Gly Asn Ser Glu Glu Lys
                440                 445                 450

Leu Val Ser Lys Thr Thr Lys Met Leu Ser Asp Pro Met Ser Gln
                455                 460                 465

Ser Val Ala Asp Leu Pro Pro Lys Leu Gln Lys Met Ala Gly Gly
                470                 475                 480

Ala Ala Arg Met Glu Gly Asn Leu Pro Ala Lys Leu Arg Lys Met
                485                 490                 495
```

```
 Asn Ser Asp Arg Phe Thr
                 500

<210> 129
<211> 1148
<212> PRT
<213> Homo sapiens

<400> 129
 Met Pro Leu Phe Lys Asn Thr Ser Val Ser Ser Leu Tyr Ser Gly
   1               5                  10                  15

 Cys Arg Leu Thr Leu Leu Arg Pro Glu Lys Asp Gly Ala Ala Thr
                  20                  25                  30

 Arg Val Asp Ala Val Cys Thr His Arg Pro Asp Pro Lys Ser Pro
                  35                  40                  45

 Gly Leu Asp Arg Glu Arg Leu Tyr Trp Lys Leu Ser Gln Leu Thr
                  50                  55                  60

 His Gly Ile Thr Glu Leu Gly Pro Tyr Thr Leu Asp Arg His Ser
                  65                  70                  75

 Leu Tyr Val Asn Gly Phe Thr His Gln Ser Ser Met Thr Thr Thr
                  80                  85                  90

 Arg Thr Pro Asp Thr Ser Thr Met His Leu Ala Thr Ser Arg Thr
                  95                 100                 105

 Pro Ala Ser Leu Ser Gly Pro Thr Thr Ala Ser Pro Leu Leu Val
                 110                 115                 120

 Leu Phe Thr Ile Asn Phe Thr Ile Thr Asn Leu Arg Tyr Glu Glu
                 125                 130                 135

 Asn Met His His Pro Gly Ser Arg Lys Phe Asn Thr Thr Glu Arg
                 140                 145                 150

 Val Leu Gln Gly Leu Leu Arg Pro Val Phe Lys Asn Thr Ser Val
                 155                 160                 165

 Gly Pro Leu Tyr Ser Gly Cys Arg Leu Thr Leu Leu Arg Pro Lys
                 170                 175                 180

 Lys Asp Gly Ala Ala Thr Lys Val Asp Ala Ile Cys Thr Tyr Arg
                 185                 190                 195

 Pro Asp Pro Lys Ser Pro Gly Leu Asp Arg Glu Gln Leu Tyr Trp
                 200                 205                 210

 Glu Leu Ser Gln Leu Thr His Ser Ile Thr Glu Leu Gly Pro Tyr
                 215                 220                 225

 Thr Leu Asp Arg Asp Ser Leu Tyr Val Asn Gly Phe Thr Gln Arg
                 230                 235                 240

 Ser Ser Val Pro Thr Thr Ser Ile Pro Gly Thr Pro Thr Val Asp
                 245                 250                 255
```

```
Leu Gly Thr Ser Gly Thr Pro Val Ser Lys Pro Gly Pro Ser Ala
                260                 265                 270

Ala Ser Pro Leu Leu Val Leu Phe Thr Leu Asn Phe Thr Ile Thr
                275                 280                 285

Asn Leu Arg Tyr Glu Glu Asn Met Gln His Pro Gly Ser Arg Lys
                290                 295                 300

Phe Asn Thr Thr Glu Arg Val Leu Gln Gly Leu Leu Arg Ser Leu
                305                 310                 315

Phe Lys Ser Thr Ser Val Gly Pro Leu Tyr Ser Gly Cys Arg Leu
                320                 325                 330

Thr Leu Leu Arg Pro Glu Lys Asp Gly Thr Ala Thr Gly Val Asp
                335                 340                 345

Ala Ile Cys Thr His His Pro Asp Pro Lys Ser Pro Arg Leu Asp
                350                 355                 360

Arg Glu Gln Leu Tyr Trp Glu Leu Ser Gln Leu Thr His Asn Ile
                365                 370                 375

Thr Glu Leu Gly His Tyr Ala Leu Asp Asn Asp Ser Leu Phe Val
                380                 385                 390

Asn Gly Phe Thr His Arg Ser Ser Val Ser Thr Thr Ser Thr Pro
                395                 400                 405

Gly Thr Pro Thr Val Tyr Leu Gly Ala Ser Lys Thr Pro Ala Ser
                410                 415                 420

Ile Phe Gly Pro Ser Ala Ala Ser His Leu Leu Ile Leu Phe Thr
                425                 430                 435

Leu Asn Phe Thr Ile Thr Asn Leu Arg Tyr Glu Glu Asn Met Trp
                440                 445                 450

Pro Gly Ser Arg Lys Phe Asn Thr Thr Glu Arg Val Leu Gln Gly
                455                 460                 465

Leu Leu Arg Pro Leu Phe Lys Asn Thr Ser Val Gly Pro Leu Tyr
                470                 475                 480

Ser Gly Ser Arg Leu Thr Leu Leu Arg Pro Glu Lys Asp Gly Glu
                485                 490                 495

Ala Thr Gly Val Asp Ala Ile Cys Thr His Arg Pro Asp Pro Thr
                500                 505                 510

Gly Pro Gly Leu Asp Arg Glu Gln Leu Tyr Leu Glu Leu Ser Gln
                515                 520                 525

Leu Thr His Ser Ile Thr Glu Leu Gly Pro Tyr Thr Leu Asp Arg
                530                 535                 540

Asp Ser Leu Tyr Val Asn Gly Phe Thr His Arg Ser Ser Val Pro
                545                 550                 555

Thr Thr Ser Thr Gly Val Val Ser Glu Glu Pro Phe Thr Leu Asn
```

```
                560                 565                 570

Phe Thr Ile Asn Asn Leu Arg Tyr Met Ala Asp Met Gly Gln Pro
                575                 580                 585

Gly Ser Leu Lys Phe Asn Ile Thr Asp Asn Val Met Lys His Leu
                590                 595                 600

Leu Ser Pro Leu Phe Gln Arg Ser Ser Leu Gly Ala Arg Tyr Thr
                605                 610                 615

Gly Cys Arg Val Ile Ala Leu Arg Ser Val Lys Asn Gly Ala Glu
                620                 625                 630

Thr Arg Val Asp Leu Leu Cys Thr Tyr Leu Gln Pro Leu Ser Gly
                635                 640                 645

Pro Gly Leu Pro Ile Lys Gln Val Phe His Glu Leu Ser Gln Gln
                650                 655                 660

Thr His Gly Ile Thr Arg Leu Gly Pro Tyr Ser Leu Asp Lys Asp
                665                 670                 675

Ser Leu Tyr Leu Asn Gly Tyr Asn Glu Pro Gly Leu Asp Glu Pro
                680                 685                 690

Pro Thr Thr Pro Lys Pro Ala Thr Thr Phe Leu Pro Pro Leu Ser
                695                 700                 705

Glu Ala Thr Thr Ala Met Gly Tyr His Leu Lys Thr Leu Thr Leu
                710                 715                 720

Asn Phe Thr Ile Ser Asn Leu Gln Tyr Ser Pro Asp Met Gly Lys
                725                 730                 735

Gly Ser Ala Thr Phe Asn Ser Thr Glu Gly Val Leu Gln His Leu
                740                 745                 750

Leu Arg Pro Leu Phe Gln Lys Ser Ser Met Gly Pro Phe Tyr Leu
                755                 760                 765

Gly Cys Gln Leu Ile Ser Leu Arg Pro Glu Lys Asp Gly Ala Ala
                770                 775                 780

Thr Gly Val Asp Thr Thr Cys Thr Tyr His Pro Asp Pro Val Gly
                785                 790                 795

Pro Gly Leu Asp Ile Gln Gln Leu Tyr Trp Glu Leu Ser Gln Leu
                800                 805                 810

Thr His Gly Val Thr Gln Leu Gly Phe Tyr Val Leu Asp Arg Asp
                815                 820                 825

Ser Leu Phe Ile Asn Gly Tyr Ala Pro Gln Asn Leu Ser Ile Arg
                830                 835                 840

Gly Glu Tyr Gln Ile Asn Phe His Ile Val Asn Trp Asn Leu Ser
                845                 850                 855

Asn Pro Asp Pro Thr Ser Ser Glu Tyr Ile Thr Leu Leu Arg Asp
                860                 865                 870
```

```
Ile Gln Asp Lys Val Thr Thr Leu Tyr Lys Gly Ser Gln Leu His
                875                 880                 885

Asp Thr Phe Arg Phe Cys Leu Val Thr Asn Leu Thr Met Asp Ser
                890                 895                 900

Val Leu Val Thr Val Lys Ala Leu Phe Ser Ser Asn Leu Asp Pro
                905                 910                 915

Ser Leu Val Glu Gln Val Phe Leu Asp Lys Thr Leu Asn Ala Ser
                920                 925                 930

Phe His Trp Leu Gly Ser Thr Tyr Gln Leu Val Asp Ile His Val
                935                 940                 945

Thr Glu Met Glu Ser Ser Val Tyr Gln Pro Thr Ser Ser Ser Ser
                950                 955                 960

Thr Gln His Phe Tyr Pro Asn Phe Thr Ile Thr Asn Leu Pro Tyr
                965                 970                 975

Ser Gln Asp Lys Ala Gln Pro Gly Thr Thr Asn Tyr Gln Arg Asn
                980                 985                 990

Lys Arg Asn Ile Glu Asp Ala Leu Asn Gln Leu Phe Arg Asn Ser
                995                1000                1005

Ser Ile Lys Ser Tyr Phe Ser Asp Cys Gln Val Ser Thr Phe Arg
               1010                1015                1020

Ser Val Pro Asn Arg His His Thr Gly Val Asp Ser Leu Cys Asn
               1025                1030                1035

Phe Ser Pro Leu Ala Arg Arg Val Asp Arg Val Ala Ile Tyr Glu
               1040                1045                1050

Glu Phe Leu Arg Met Thr Arg Asn Gly Thr Gln Leu Gln Asn Phe
               1055                1060                1065

Thr Leu Asp Arg Ser Ser Val Leu Val Asp Gly Tyr Ser Pro Asn
               1070                1075                1080

Arg Asn Glu Pro Leu Thr Gly Asn Ser Asp Leu Pro Phe Trp Ala
               1085                1090                1095

Val Ile Phe Ile Gly Leu Ala Gly Leu Leu Gly Leu Ile Thr Cys
               1100                1105                1110

Leu Ile Cys Gly Val Leu Val Thr Thr Arg Arg Arg Lys Lys Glu
               1115                1120                1125

Gly Glu Tyr Asn Val Gln Gln Gln Cys Pro Gly Tyr Tyr Gln Ser
               1130                1135                1140

His Leu Asp Leu Glu Asp Leu Gln
               1145
```

<210> 130
<211> 564
<212> PRT

<213> Homo sapiens

<400> 130

```
 Met Gly Asn Trp Val Val Asn His Trp Phe Ser Val Leu Phe Leu
   1               5                  10                  15

 Val Val Trp Leu Gly Leu Asn Val Phe Leu Phe Val Asp Ala Phe
                20                  25                  30

 Leu Lys Tyr Glu Lys Ala Asp Lys Tyr Tyr Tyr Thr Arg Lys Ile
                35                  40                  45

 Leu Gly Ser Thr Leu Ala Cys Ala Arg Ala Ser Ala Leu Cys Leu
                50                  55                  60

 Asn Phe Asn Ser Thr Leu Ile Leu Leu Pro Val Cys Arg Asn Leu
                65                  70                  75

 Leu Ser Phe Leu Arg Gly Thr Cys Ser Phe Cys Ser Arg Thr Leu
                80                  85                  90

 Arg Lys Gln Leu Asp His Asn Leu Thr Phe His Lys Leu Val Ala
                95                 100                 105

 Tyr Met Ile Cys Leu His Thr Ala Ile His Ile Ile Ala His Leu
               110                 115                 120

 Phe Asn Phe Asp Cys Tyr Ser Arg Ser Arg Gln Ala Thr Asp Gly
               125                 130                 135

 Ser Leu Ala Ser Ile Leu Ser Ser Leu Ser His Asp Glu Lys Lys
               140                 145                 150

 Gly Gly Ser Trp Leu Asn Pro Ile Gln Ser Arg Asn Thr Thr Val
               155                 160                 165

 Glu Tyr Val Thr Phe Thr Ser Val Ala Gly Leu Thr Gly Val Ile
               170                 175                 180

 Met Thr Ile Ala Leu Ile Leu Met Val Thr Ser Ala Thr Glu Phe
               185                 190                 195

 Ile Arg Arg Ser Tyr Phe Glu Val Phe Trp Tyr Thr His His Leu
               200                 205                 210

 Phe Ile Phe Tyr Ile Leu Gly Leu Gly Ile His Gly Ile Gly Gly
               215                 220                 225

 Ile Val Arg Gly Gln Thr Glu Glu Ser Met Asn Glu Ser His Pro
               230                 235                 240

 Arg Lys Cys Ala Glu Ser Phe Glu Met Trp Asp Asp Arg Asp Ser
               245                 250                 255

 His Cys Arg Arg Pro Lys Phe Glu Gly His Pro Pro Glu Ser Trp
               260                 265                 270

 Lys Trp Ile Leu Ala Pro Val Ile Leu Tyr Ile Cys Glu Arg Ile
               275                 280                 285

 Leu Arg Phe Tyr Arg Ser Gln Gln Lys Val Val Ile Thr Lys Val
```

```
                290                 295                 300

 Val Met His Pro Ser Lys Val Leu Glu Leu Gln Met Asn Lys Arg
                305                 310                 315

 Gly Phe Ser Met Glu Val Gly Gln Tyr Ile Phe Val Asn Cys Pro
                320                 325                 330

 Ser Ile Ser Leu Leu Glu Trp His Pro Phe Thr Leu Thr Ser Ala
                335                 340                 345

 Pro Glu Glu Asp Phe Phe Ser Ile His Ile Arg Ala Ala Gly Asp
                350                 355                 360

 Trp Thr Glu Asn Leu Ile Arg Ala Phe Glu Gln Gln Tyr Ser Pro
                365                 370                 375

 Ile Pro Arg Ile Glu Val Asp Gly Pro Phe Gly Thr Ala Ser Glu
                380                 385                 390

 Asp Val Phe Gln Tyr Glu Val Ala Val Leu Val Gly Ala Gly Ile
                395                 400                 405

 Gly Val Thr Pro Phe Ala Ser Ile Leu Lys Ser Ile Trp Tyr Lys
                410                 415                 420

 Phe Gln Cys Ala Asp His Asn Leu Lys Thr Lys Lys Ile Tyr Phe
                425                 430                 435

 Tyr Trp Ile Cys Arg Glu Thr Gly Ala Phe Ser Trp Phe Asn Asn
                440                 445                 450

 Leu Leu Thr Ser Leu Glu Gln Glu Met Glu Glu Leu Gly Lys Val
                455                 460                 465

 Gly Phe Leu Asn Tyr Arg Leu Phe Leu Thr Gly Trp Asp Ser Asn
                470                 475                 480

 Ile Val Gly His Ala Ala Leu Asn Phe Asp Lys Ala Thr Asp Ile
                485                 490                 495

 Val Thr Gly Leu Lys Gln Lys Thr Ser Phe Gly Arg Pro Met Trp
                500                 505                 510

 Asp Asn Glu Phe Ser Thr Ile Ala Thr Ser His Pro Lys Ser Val
                515                 520                 525

 Val Gly Val Phe Leu Cys Gly Pro Arg Thr Leu Ala Lys Ser Leu
                530                 535                 540

 Arg Lys Cys Cys His Arg Tyr Ser Ser Leu Asp Pro Arg Lys Val
                545                 550                 555

 Gln Phe Tyr Phe Asn Lys Glu Asn Phe
                560

<210> 131
<211> 564
<212> PRT
<213> Homo sapiens
```

<400> 131

```
Met Gly Asn Trp Val Val Asn His Trp Phe Ser Val Leu Phe Leu
  1               5                  10                  15

Val Val Trp Leu Gly Leu Asn Val Phe Leu Phe Val Asp Ala Phe
                20                  25                  30

Leu Lys Tyr Glu Lys Ala Asp Lys Tyr Tyr Tyr Thr Arg Lys Ile
                35                  40                  45

Leu Gly Ser Thr Leu Ala Cys Ala Arg Ala Ser Ala Leu Cys Leu
                50                  55                  60

Asn Phe Asn Ser Thr Leu Ile Leu Leu Pro Val Cys Arg Asn Leu
                65                  70                  75

Leu Ser Phe Leu Arg Gly Thr Cys Ser Phe Cys Ser Arg Thr Leu
                80                  85                  90

Arg Lys Gln Leu Asp His Asn Leu Thr Phe His Lys Leu Val Ala
                95                 100                 105

Tyr Met Ile Cys Leu His Thr Ala Ile His Ile Ile Ala His Leu
               110                 115                 120

Phe Asn Phe Asp Cys Tyr Ser Arg Ser Arg Gln Ala Thr Asp Gly
               125                 130                 135

Ser Leu Ala Ser Ile Leu Ser Ser Leu Ser His Asp Glu Lys Lys
               140                 145                 150

Gly Gly Ser Trp Leu Asn Pro Ile Gln Ser Arg Asn Thr Thr Val
               155                 160                 165

Glu Tyr Val Thr Phe Thr Ser Ile Ala Gly Leu Thr Gly Val Ile
               170                 175                 180

Met Thr Ile Ala Leu Ile Leu Met Val Thr Ser Ala Thr Glu Phe
               185                 190                 195

Ile Arg Arg Ser Tyr Phe Glu Val Phe Trp Tyr Thr His His Leu
               200                 205                 210

Phe Ile Phe Tyr Ile Leu Gly Leu Gly Ile His Gly Ile Gly Gly
               215                 220                 225

Ile Val Arg Gly Gln Thr Glu Glu Ser Met Asn Glu Ser His Pro
               230                 235                 240

Arg Lys Cys Ala Glu Ser Phe Glu Met Trp Asp Asp Arg Asp Ser
               245                 250                 255

His Cys Arg Arg Pro Lys Phe Glu Gly His Pro Pro Glu Ser Trp
               260                 265                 270

Lys Trp Ile Leu Ala Pro Val Ile Leu Tyr Ile Cys Glu Arg Ile
               275                 280                 285

Leu Arg Phe Tyr Arg Ser Gln Gln Lys Val Val Ile Thr Lys Val
               290                 295                 300
```

Val Met His Pro Ser Lys Val Leu Glu Leu Gln Met Asn Lys Arg
305 310 315

Gly Phe Ser Met Glu Val Gly Gln Tyr Ile Phe Val Asn Cys Pro
320 325 330

Ser Ile Ser Leu Leu Glu Trp His Pro Phe Thr Leu Thr Ser Ala
335 340 345

Pro Glu Glu Asp Phe Phe Ser Ile His Ile Arg Ala Ala Gly Asp
350 355 360

Trp Thr Glu Asn Leu Ile Arg Ala Phe Glu Gln Gln Tyr Ser Pro
365 370 375

Ile Pro Arg Ile Glu Val Asp Gly Pro Phe Gly Thr Ala Ser Glu
380 385 390

Asp Val Phe Gln Tyr Glu Val Ala Val Leu Val Gly Ala Gly Ile
395 400 405

Gly Val Thr Pro Phe Ala Ser Ile Leu Lys Ser Ile Trp Tyr Lys
410 415 420

Phe Gln Cys Ala Asp His Asn Leu Lys Thr Lys Lys Ile Tyr Phe
425 430 435

Tyr Trp Ile Cys Arg Glu Thr Gly Ala Phe Ser Trp Phe Asn Asn
440 445 450

Leu Leu Thr Ser Leu Glu Gln Glu Met Glu Glu Leu Gly Lys Val
455 460 465

Gly Phe Leu Asn Tyr Arg Leu Phe Leu Thr Gly Trp Asp Ser Asn
470 475 480

Ile Val Gly His Ala Ala Leu Asn Phe Asp Lys Ala Thr Asp Ile
485 490 495

Val Thr Gly Leu Lys Gln Lys Thr Ser Phe Gly Arg Pro Met Trp
500 505 510

Asp Asn Glu Phe Ser Thr Ile Ala Thr Ser His Pro Lys Ser Val
515 520 525

Val Gly Val Phe Leu Cys Gly Pro Arg Thr Leu Ala Lys Ser Leu
530 535 540

Arg Lys Cys Cys His Arg Tyr Ser Ser Leu Asp Pro Arg Lys Val
545 550 555

Gln Phe Tyr Phe Asn Lys Glu Asn Phe
560

<210> 132
<211> 294
<212> PRT
<213> Homo sapiens

<400> 132
Met Gly Asn Trp Val Val Asn His Trp Phe Ser Val Leu Phe Leu

```
  1               5                  10                  15

Val Val Trp Leu Gly Leu Asn Val Phe Leu Phe Val Asp Ala Phe
                20                  25                  30

Leu Lys Tyr Glu Lys Ala Asp Lys Tyr Tyr Tyr Thr Arg Lys Ile
                35                  40                  45

Leu Gly Ser Trp Lys Trp Ile Leu Ala Pro Val Ile Leu Tyr Ile
                50                  55                  60

Cys Glu Arg Ile Leu Arg Phe Tyr Arg Ser Gln Gln Lys Val Val
                65                  70                  75

Ile Thr Lys Val Val Met His Pro Ser Lys Val Leu Glu Leu Gln
                80                  85                  90

Met Asn Lys Arg Gly Phe Ser Met Glu Val Gly Gln Tyr Ile Phe
                95                 100                 105

Val Asn Cys Pro Ser Ile Ser Leu Leu Glu Trp His Pro Phe Thr
               110                 115                 120

Leu Thr Ser Ala Pro Glu Glu Asp Phe Phe Ser Ile His Ile Arg
               125                 130                 135

Ala Ala Gly Asp Trp Thr Glu Asn Leu Ile Arg Ala Phe Glu Gln
               140                 145                 150

Gln Tyr Ser Pro Ile Pro Arg Ile Glu Val Asp Gly Pro Phe Gly
               155                 160                 165

Thr Ala Ser Glu Asp Val Phe Gln Tyr Glu Val Ala Val Leu Val
               170                 175                 180

Gly Ala Gly Ile Gly Val Thr Pro Phe Ala Ser Ile Leu Lys Ser
               185                 190                 195

Ile Trp Tyr Lys Phe Gln Cys Ala Asp His Asn Leu Lys Thr Lys
               200                 205                 210

Lys Val Gly His Ala Ala Leu Asn Phe Asp Lys Ala Thr Asp Ile
               215                 220                 225

Val Thr Gly Leu Lys Gln Lys Thr Ser Phe Gly Arg Pro Met Trp
               230                 235                 240

Asp Asn Glu Phe Ser Thr Ile Ala Thr Ser His Pro Lys Ser Val
               245                 250                 255

Val Gly Val Phe Leu Cys Gly Pro Arg Thr Leu Ala Lys Ser Leu
               260                 265                 270

Arg Lys Cys Cys His Arg Tyr Ser Ser Leu Asp Pro Arg Lys Val
               275                 280                 285

Gln Phe Tyr Phe Asn Lys Glu Asn Phe
               290
```

<210> 133
<211> 515

<212> PRT
<213> Homo sapiens

<400> 133
Met Gly Asn Trp Val Val Asn His Trp Phe Ser Val Leu Phe Leu
1 5 10 15

Val Val Trp Leu Gly Leu Asn Val Phe Leu Phe Val Asp Ala Phe
20 25 30

Leu Lys Tyr Glu Lys Ala Asp Lys Tyr Tyr Tyr Thr Arg Lys Ile
35 40 45

Leu Gly Ser Thr Leu Ala Cys Ala Arg Ala Ser Ala Leu Cys Leu
50 55 60

Asn Phe Asn Ser Thr Leu Ile Leu Leu Pro Val Cys Arg Asn Leu
65 70 75

Leu Ser Phe Leu Arg Gly Thr Cys Ser Phe Cys Ser Arg Thr Leu
80 85 90

Arg Lys Gln Leu Asp His Asn Leu Thr Phe His Lys Leu Val Ala
95 100 105

Tyr Met Ile Cys Leu His Thr Ala Ile His Ile Ile Ala His Leu
110 115 120

Phe Asn Phe Asp Cys Tyr Ser Arg Ser Arg Gln Ala Thr Asp Gly
125 130 135

Ser Leu Ala Ser Ile Leu Ser Ser Leu Ser His Asp Glu Lys Lys
140 145 150

Gly Gly Ser Trp Leu Asn Pro Ile Gln Ser Arg Asn Thr Thr Val
155 160 165

Glu Tyr Val Thr Phe Thr Ser Ile Ala Gly Leu Thr Gly Val Ile
170 175 180

Met Thr Ile Ala Leu Ile Leu Met Val Thr Ser Ala Thr Glu Phe
185 190 195

Ile Arg Arg Ser Tyr Phe Glu Val Phe Trp Tyr Thr His His Leu
200 205 210

Phe Ile Phe Tyr Ile Leu Gly Leu Gly Ile His Gly Ile Gly Gly
215 220 225

Ile Val Arg Gly Gln Thr Glu Glu Ser Met Asn Glu Ser His Pro
230 235 240

Arg Lys Cys Ala Glu Ser Phe Glu Met Trp Asp Asp Arg Asp Ser
245 250 255

His Cys Arg Arg Pro Lys Phe Glu Gly His Pro Pro Glu Ser Trp
260 265 270

Lys Trp Ile Leu Ala Pro Val Ile Leu Tyr Ile Cys Glu Arg Ile
275 280 285

```
 Leu Arg Phe Tyr Arg Ser Gln Gln Lys Val Val Ile Thr Lys Val
                 290                 295                 300

 Val Met His Pro Ser Lys Val Leu Glu Leu Gln Met Asn Lys Arg
                 305                 310                 315

 Gly Phe Ser Met Glu Val Gly Gln Tyr Ile Phe Val Asn Cys Pro
                 320                 325                 330

 Ser Ile Ser Leu Leu Glu Trp His Pro Phe Thr Leu Thr Ser Ala
                 335                 340                 345

 Pro Glu Glu Asp Phe Phe Ser Ile His Ile Arg Ala Ala Gly Asp
                 350                 355                 360

 Trp Thr Glu Asn Leu Ile Arg Ala Phe Glu Gln Gln Tyr Ser Pro
                 365                 370                 375

 Ile Pro Arg Ile Glu Val Asp Gly Pro Phe Gly Thr Ala Ser Glu
                 380                 385                 390

 Asp Val Phe Gln Tyr Glu Val Ala Val Leu Val Gly Ala Gly Ile
                 395                 400                 405

 Gly Val Thr Pro Phe Ala Ser Ile Leu Lys Ser Ile Trp Tyr Lys
                 410                 415                 420

 Phe Gln Cys Ala Asp His Asn Leu Lys Thr Lys Lys Val Gly His
                 425                 430                 435

 Ala Ala Leu Asn Phe Asp Lys Ala Thr Asp Ile Val Thr Gly Leu
                 440                 445                 450

 Lys Gln Lys Thr Ser Phe Gly Arg Pro Met Trp Asp Asn Glu Phe
                 455                 460                 465

 Ser Thr Ile Ala Thr Ser His Pro Lys Ser Val Val Gly Val Phe
                 470                 475                 480

 Leu Cys Gly Pro Arg Thr Leu Ala Lys Ser Leu Arg Lys Cys Cys
                 485                 490                 495

 His Arg Tyr Ser Ser Leu Asp Pro Arg Lys Val Gln Phe Tyr Phe
                 500                 505                 510

 Asn Lys Glu Asn Phe
                 515

<210> 134
<211> 757
<212> PRT
<213> Homo sapiens

<400> 134
 Met Thr Ser Gly Ser Val Phe Phe Tyr Ile Leu Ile Phe Gly Lys
   1               5                  10                  15

 Tyr Phe Ser His Gly Gly Gly Gln Asp Val Lys Cys Ser Leu Gly
                  20                  25                  30

 Tyr Phe Pro Cys Gly Asn Ile Thr Lys Cys Leu Pro Gln Leu Leu
```

```
                35                  40                  45

His Cys Asn Gly Val Asp Asp Cys Gly Asn Gln Ala Asp Glu Asp
                50                  55                  60

Asn Cys Gly Asp Asn Asn Gly Trp Ser Met Gln Phe Asp Lys Tyr
                65                  70                  75

Phe Ala Ser Tyr Tyr Lys Met Thr Ser Gln Tyr Pro Phe Glu Ala
                80                  85                  90

Glu Thr Pro Glu Cys Leu Val Gly Ser Val Pro Val Gln Cys Leu
                95                 100                 105

Cys Gln Gly Leu Glu Leu Asp Cys Asp Glu Thr Asn Leu Arg Ala
               110                 115                 120

Val Pro Ser Val Ser Ser Asn Val Thr Ala Met Ser Leu Gln Trp
               125                 130                 135

Asn Leu Ile Arg Lys Leu Pro Pro Asp Cys Phe Lys Asn Tyr His
               140                 145                 150

Asp Leu Gln Lys Leu Tyr Leu Gln Asn Asn Lys Ile Thr Ser Ile
               155                 160                 165

Ser Ile Tyr Ala Phe Arg Gly Leu Asn Ser Leu Thr Lys Leu Tyr
               170                 175                 180

Leu Ser His Asn Arg Ile Thr Phe Leu Lys Pro Gly Val Phe Glu
               185                 190                 195

Asp Leu His Arg Leu Glu Trp Leu Ile Ile Glu Asp Asn His Leu
               200                 205                 210

Ser Arg Ile Ser Pro Pro Thr Phe Tyr Gly Leu Asn Ser Leu Ile
               215                 220                 225

Leu Leu Val Leu Met Asn Asn Val Leu Thr Arg Leu Pro Asp Lys
               230                 235                 240

Pro Leu Cys Gln His Met Pro Arg Leu His Trp Leu Asp Leu Glu
               245                 250                 255

Gly Asn His Ile His Asn Leu Arg Asn Leu Thr Phe Ile Ser Cys
               260                 265                 270

Ser Asn Leu Thr Val Leu Val Met Arg Lys Asn Lys Ile Asn His
               275                 280                 285

Leu Asn Glu Asn Thr Phe Ala Pro Leu Gln Lys Leu Asp Glu Leu
               290                 295                 300

Asp Leu Gly Ser Asn Lys Ile Glu Asn Leu Pro Pro Leu Ile Phe
               305                 310                 315

Lys Asp Leu Lys Glu Leu Ser Gln Leu Asn Leu Ser Tyr Asn Pro
               320                 325                 330

Ile Gln Lys Ile Gln Ala Asn Gln Phe Asp Tyr Leu Val Lys Leu
               335                 340                 345
```

```
Lys Ser Leu Ser Leu Glu Gly Ile Glu Ile Ser Asn Ile Gln Gln
                350                 355                 360

Arg Met Phe Arg Pro Leu Met Asn Leu Ser His Ile Tyr Phe Lys
                365                 370                 375

Lys Phe Gln Tyr Cys Gly Tyr Ala Pro His Val Arg Ser Cys Lys
                380                 385                 390

Pro Asn Thr Asp Gly Ile Ser Ser Leu Glu Asn Leu Leu Ala Ser
                395                 400                 405

Ile Ile Gln Arg Val Phe Val Trp Val Val Ser Ala Val Thr Cys
                410                 415                 420

Phe Gly Asn Ile Phe Val Ile Cys Met Arg Pro Tyr Ile Arg Ser
                425                 430                 435

Glu Asn Lys Leu Tyr Ala Met Ser Ile Ile Ser Leu Cys Cys Ala
                440                 445                 450

Asp Cys Leu Met Gly Ile Tyr Leu Phe Val Ile Gly Gly Phe Asp
                455                 460                 465

Leu Lys Phe Arg Gly Glu Tyr Asn Lys His Ala Gln Leu Trp Met
                470                 475                 480

Glu Ser Thr His Cys Gln Leu Val Gly Ser Leu Ala Ile Leu Ser
                485                 490                 495

Thr Glu Val Ser Val Leu Leu Leu Thr Phe Leu Thr Leu Glu Lys
                500                 505                 510

Tyr Ile Cys Ile Val Tyr Pro Phe Arg Cys Val Arg Pro Gly Lys
                515                 520                 525

Cys Arg Thr Ile Thr Val Leu Ile Leu Ile Trp Ile Thr Gly Phe
                530                 535                 540

Ile Val Ala Phe Ile Pro Leu Ser Asn Lys Glu Phe Phe Lys Asn
                545                 550                 555

Tyr Tyr Gly Thr Asn Gly Val Cys Phe Pro Leu His Ser Glu Asp
                560                 565                 570

Thr Glu Ser Ile Gly Ala Gln Ile Tyr Ser Val Ala Ile Phe Leu
                575                 580                 585

Gly Ile Asn Leu Ala Ala Phe Ile Ile Ile Val Phe Ser Tyr Gly
                590                 595                 600

Ser Met Phe Tyr Ser Val His Gln Ser Ala Ile Thr Ala Thr Glu
                605                 610                 615

Ile Arg Asn Gln Val Lys Lys Glu Met Ile Leu Ala Lys Arg Phe
                620                 625                 630

Phe Phe Ile Val Phe Thr Asp Ala Leu Cys Trp Ile Pro Ile Phe
                635                 640                 645
```

```
 Val Val Lys Phe Leu Ser Leu Leu Gln Val Glu Ile Pro Gly Thr
                 650                 655                 660

 Ile Thr Ser Trp Val Val Ile Phe Ile Leu Pro Ile Asn Ser Ala
                 665                 670                 675

 Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Arg Pro Phe Lys Glu
                 680                 685                 690

 Met Ile His Arg Phe Trp Tyr Asn Tyr Arg Gln Arg Lys Ser Met
                 695                 700                 705

 Asp Ser Lys Gly Gln Lys Thr Tyr Ala Pro Ser Phe Ile Trp Val
                 710                 715                 720

 Glu Met Trp Pro Leu Gln Glu Met Pro Pro Glu Leu Met Lys Pro
                 725                 730                 735

 Asp Leu Phe Thr Tyr Pro Cys Glu Met Ser Leu Ile Ser Gln Ser
                 740                 745                 750

 Thr Arg Leu Asn Ser Tyr Ser
                 755

<210> 135
<211> 757
<212> PRT
<213> Homo sapiens

<400> 135
 Met Thr Ser Gly Ser Val Phe Phe Tyr Ile Leu Ile Phe Gly Lys
   1               5                  10                  15

 Tyr Phe Ser His Gly Gly Gly Gln Asp Val Lys Cys Ser Leu Gly
                  20                  25                  30

 Tyr Phe Pro Cys Gly Asn Ile Thr Lys Cys Leu Pro Gln Leu Leu
                  35                  40                  45

 His Cys Asn Gly Val Asp Asp Cys Gly Asn Gln Ala Asp Glu Asp
                  50                  55                  60

 Asn Cys Gly Asp Asn Asn Gly Trp Ser Leu Gln Phe Asp Lys Tyr
                  65                  70                  75

 Phe Ala Ser Tyr Tyr Lys Met Thr Ser Gln Tyr Pro Phe Glu Ala
                  80                  85                  90

 Glu Thr Pro Glu Cys Leu Val Gly Ser Val Pro Val Gln Cys Leu
                  95                 100                 105

 Cys Gln Gly Leu Glu Leu Asp Cys Asp Glu Thr Asn Leu Arg Ala
                 110                 115                 120

 Val Pro Ser Val Ser Ser Asn Val Thr Ala Met Ser Leu Gln Trp
                 125                 130                 135

 Asn Leu Ile Arg Lys Leu Pro Pro Asp Cys Phe Lys Asn Tyr His
                 140                 145                 150

 Asp Leu Gln Lys Leu Tyr Leu Gln Asn Asn Lys Ile Thr Ser Ile
```

```
                155                 160                 165

Ser Ile Tyr Ala Phe Arg Gly Leu Asn Ser Leu Thr Lys Leu Tyr
                170                 175                 180

Leu Ser His Asn Arg Ile Thr Phe Leu Lys Pro Gly Val Phe Glu
                185                 190                 195

Asp Leu His Arg Leu Glu Trp Leu Ile Ile Glu Asp Asn His Leu
                200                 205                 210

Ser Arg Ile Ser Pro Pro Thr Phe Tyr Gly Leu Asn Ser Leu Ile
                215                 220                 225

Leu Leu Val Leu Met Asn Asn Val Leu Thr Arg Leu Pro Asp Lys
                230                 235                 240

Pro Leu Cys Gln His Met Pro Arg Leu His Trp Leu Asp Leu Glu
                245                 250                 255

Gly Asn His Ile His Asn Leu Arg Asn Leu Thr Phe Ile Ser Cys
                260                 265                 270

Ser Asn Leu Thr Val Leu Val Met Arg Lys Asn Lys Ile Asn His
                275                 280                 285

Leu Asn Glu Asn Thr Phe Ala Pro Leu Gln Lys Leu Asp Glu Leu
                290                 295                 300

Asp Leu Gly Ser Asn Lys Ile Glu Asn Leu Pro Pro Leu Ile Phe
                305                 310                 315

Lys Asp Leu Lys Glu Leu Ser Gln Leu Asn Leu Ser Tyr Asn Pro
                320                 325                 330

Ile Gln Lys Ile Gln Ala Asn Gln Phe Asp Tyr Leu Val Lys Leu
                335                 340                 345

Lys Ser Leu Ser Leu Glu Gly Ile Glu Ile Ser Asn Ile Gln Gln
                350                 355                 360

Arg Met Phe Arg Pro Leu Met Asn Leu Ser His Ile Tyr Phe Lys
                365                 370                 375

Lys Phe Gln Tyr Cys Gly Tyr Ala Pro His Val Arg Ser Cys Lys
                380                 385                 390

Pro Asn Thr Asp Gly Ile Ser Ser Leu Glu Asn Leu Leu Ala Ser
                395                 400                 405

Ile Ile Gln Arg Val Phe Val Trp Val Val Ser Ala Val Thr Cys
                410                 415                 420

Phe Gly Asn Ile Phe Val Ile Cys Met Arg Pro Tyr Ile Arg Ser
                425                 430                 435

Glu Asn Lys Leu Tyr Ala Met Ser Ile Ile Ser Leu Cys Cys Ala
                440                 445                 450

Asp Cys Leu Met Gly Ile Tyr Leu Phe Val Ile Gly Gly Phe Asp
                455                 460                 465
```

```
Leu Lys Phe Arg Gly Glu Tyr Asn Lys His Ala Gln Leu Trp Met
                470                 475                 480

Glu Ser Thr His Cys Gln Leu Val Gly Ser Leu Ala Ile Leu Ser
                485                 490                 495

Thr Glu Val Ser Val Leu Leu Leu Thr Phe Leu Thr Leu Glu Lys
                500                 505                 510

Tyr Ile Cys Ile Val Tyr Pro Phe Arg Cys Val Arg Pro Gly Lys
                515                 520                 525

Cys Arg Thr Ile Thr Val Leu Ile Leu Ile Trp Ile Thr Gly Phe
                530                 535                 540

Ile Val Ala Phe Ile Pro Leu Ser Asn Lys Glu Phe Phe Lys Asn
                545                 550                 555

Tyr Tyr Gly Thr Asn Gly Val Cys Phe Pro Leu His Ser Glu Asp
                560                 565                 570

Thr Glu Ser Ile Gly Ala Gln Ile Tyr Ser Val Ala Ile Phe Leu
                575                 580                 585

Gly Ile Asn Leu Ala Ala Phe Ile Ile Ile Val Phe Ser Tyr Gly
                590                 595                 600

Ser Met Phe Tyr Ser Val His Gln Ser Ala Ile Thr Ala Thr Glu
                605                 610                 615

Ile Arg Asn Gln Val Lys Lys Glu Met Ile Leu Ala Lys Arg Phe
                620                 625                 630

Phe Phe Ile Val Phe Thr Asp Ala Leu Cys Trp Ile Pro Ile Phe
                635                 640                 645

Val Val Lys Phe Leu Ser Leu Leu Gln Val Glu Ile Pro Gly Thr
                650                 655                 660

Ile Thr Ser Trp Val Val Ile Phe Ile Leu Pro Ile Asn Ser Ala
                665                 670                 675

Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Arg Pro Phe Lys Glu
                680                 685                 690

Met Ile His Arg Phe Trp Tyr Asn Tyr Arg Gln Arg Lys Ser Met
                695                 700                 705

Asp Ser Lys Gly Gln Lys Thr Tyr Ala Pro Ser Phe Ile Trp Val
                710                 715                 720

Glu Met Trp Pro Leu Gln Glu Met Pro Pro Glu Leu Met Lys Pro
                725                 730                 735

Asp Leu Phe Thr Tyr Pro Cys Glu Met Ser Leu Ile Ser Gln Ser
                740                 745                 750

Thr Arg Leu Asn Ser Tyr Ser
                755
```

<210> 136
<211> 879
<212> PRT
<213> Homo sapiens

<400> 136
Met Gly Arg Leu Ala Ser Arg Pro Leu Leu Leu Ala Leu Leu Ser
1 5 10 15

Leu Ala Leu Cys Arg Gly Arg Val Val Arg Val Pro Thr Ala Thr
20 25 30

Leu Val Arg Val Val Gly Thr Glu Leu Val Ile Pro Cys Asn Val
35 40 45

Ser Asp Tyr Asp Gly Pro Ser Glu Gln Asn Phe Asp Trp Ser Phe
50 55 60

Ser Ser Leu Gly Ser Ser Phe Val Glu Leu Ala Ser Thr Trp Glu
65 70 75

Val Gly Phe Pro Ala Gln Leu Tyr Gln Glu Arg Leu Gln Arg Gly
80 85 90

Glu Ile Leu Leu Arg Arg Thr Ala Asn Asp Ala Val Glu Leu His
95 100 105

Ile Lys Asn Val Gln Pro Ser Asp Gln Gly His Tyr Lys Cys Ser
110 115 120

Thr Pro Ser Thr Asp Ala Thr Val Gln Gly Asn Tyr Glu Asp Thr
125 130 135

Val Gln Val Lys Val Leu Ala Asp Ser Leu His Val Gly Pro Ser
140 145 150

Ala Arg Pro Pro Pro Ser Leu Ser Leu Arg Glu Gly Glu Pro Phe
155 160 165

Glu Leu Arg Cys Thr Ala Ala Ser Ala Ser Pro Leu His Thr His
170 175 180

Leu Ala Leu Leu Trp Glu Val His Arg Gly Pro Ala Arg Arg Ser
185 190 195

Val Leu Ala Leu Thr His Glu Gly Arg Phe His Pro Gly Leu Gly
200 205 210

Tyr Glu Gln Arg Tyr His Ser Gly Asp Val Arg Leu Asp Thr Val
215 220 225

Gly Ser Asp Ala Tyr Arg Leu Ser Val Ser Arg Ala Leu Ser Ala
230 235 240

Asp Gln Gly Ser Tyr Arg Cys Ile Val Ser Glu Trp Ile Ala Glu
245 250 255

Gln Gly Asn Trp Gln Glu Ile Gln Glu Lys Ala Val Glu Val Ala
260 265 270

Thr Val Val Ile Gln Pro Thr Val Leu Arg Ala Ala Val Pro Lys

```
                275                 280                 285

Asn Val Ser Val Ala Glu Gly Lys Glu Leu Asp Leu Thr Cys Asn
                290                 295                 300

Ile Thr Thr Asp Arg Ala Asp Asp Val Arg Pro Glu Val Thr Trp
                305                 310                 315

Ser Phe Ser Arg Met Pro Asp Ser Thr Leu Pro Gly Ser Arg Val
                320                 325                 330

Leu Ala Arg Leu Asp Arg Asp Ser Leu Val His Ser Ser Pro His
                335                 340                 345

Val Ala Leu Ser His Val Asp Ala Arg Ser Tyr His Leu Leu Val
                350                 355                 360

Arg Asp Val Ser Lys Glu Asn Ser Gly Tyr Tyr Tyr Cys His Val
                365                 370                 375

Ser Leu Trp Ala Pro Gly His Asn Arg Ser Trp His Lys Val Ala
                380                 385                 390

Glu Ala Val Ser Ser Pro Ala Gly Val Gly Val Thr Trp Leu Glu
                395                 400                 405

Pro Asp Tyr Gln Val Tyr Leu Asn Ala Ser Lys Val Pro Gly Phe
                410                 415                 420

Ala Asp Asp Pro Thr Glu Leu Ala Cys Arg Val Val Asp Thr Lys
                425                 430                 435

Ser Gly Glu Ala Asn Val Arg Phe Thr Val Ser Trp Tyr Tyr Arg
                440                 445                 450

Met Asn Arg Arg Ser Asp Asn Val Val Thr Ser Glu Leu Leu Ala
                455                 460                 465

Val Met Asp Gly Asp Trp Thr Leu Lys Tyr Gly Glu Arg Ser Lys
                470                 475                 480

Gln Arg Ala Gln Asp Gly Asp Phe Ile Phe Ser Lys Glu His Thr
                485                 490                 495

Asp Thr Phe Asn Phe Arg Ile Gln Arg Thr Thr Glu Glu Asp Arg
                500                 505                 510

Gly Asn Tyr Tyr Cys Val Val Ser Ala Trp Thr Lys Gln Arg Asn
                515                 520                 525

Asn Ser Trp Val Lys Ser Lys Asp Val Phe Ser Lys Pro Val Asn
                530                 535                 540

Ile Phe Trp Ala Leu Glu Asp Ser Val Leu Val Val Lys Ala Arg
                545                 550                 555

Gln Pro Lys Pro Phe Phe Ala Ala Gly Asn Thr Phe Glu Met Thr
                560                 565                 570

Cys Lys Val Ser Ser Lys Asn Ile Lys Ser Pro Arg Tyr Ser Val
                575                 580                 585
```

```
Leu Ile Met Ala Glu Lys Pro Val Gly Asp Leu Ser Ser Pro Asn
                590                 595                 600

Glu Thr Lys Tyr Ile Ile Ser Leu Asp Gln Asp Ser Val Val Lys
                605                 610                 615

Leu Glu Asn Trp Thr Asp Ala Ser Arg Val Asp Gly Val Val Leu
                620                 625                 630

Glu Lys Val Gln Glu Asp Glu Phe Arg Tyr Arg Met Tyr Gln Thr
                635                 640                 645

Gln Val Ser Asp Ala Gly Leu Tyr Arg Cys Met Val Thr Ala Trp
                650                 655                 660

Ser Pro Val Arg Gly Ser Leu Trp Arg Glu Ala Ala Thr Ser Leu
                665                 670                 675

Ser Asn Pro Ile Glu Ile Asp Phe Gln Thr Ser Gly Pro Ile Phe
                680                 685                 690

Asn Ala Ser Val His Ser Asp Thr Pro Ser Val Ile Arg Gly Asp
                695                 700                 705

Leu Ile Lys Leu Phe Cys Ile Ile Thr Val Glu Gly Ala Ala Leu
                710                 715                 720

Asp Pro Asp Asp Met Ala Phe Asp Val Ser Trp Phe Ala Val His
                725                 730                 735

Ser Phe Gly Leu Asp Lys Ala Pro Val Leu Leu Ser Ser Leu Asp
                740                 745                 750

Arg Lys Gly Ile Val Thr Thr Ser Arg Arg Asp Trp Lys Ser Asp
                755                 760                 765

Leu Ser Leu Glu Arg Val Ser Val Leu Glu Phe Leu Leu Gln Val
                770                 775                 780

His Gly Ser Glu Asp Gln Asp Phe Gly Asn Tyr Tyr Cys Ser Val
                785                 790                 795

Thr Pro Trp Val Lys Ser Pro Thr Gly Ser Trp Gln Lys Glu Ala
                800                 805                 810

Glu Ile His Ser Lys Pro Val Phe Ile Thr Val Lys Met Asp Val
                815                 820                 825

Leu Asn Ala Phe Lys Tyr Pro Leu Leu Ile Gly Ile Gly Leu Ser
                830                 835                 840

Thr Val Ile Gly Leu Leu Ser Cys Leu Ile Gly Tyr Cys Ser Ser
                845                 850                 855

His Trp Cys Cys Lys Lys Glu Val Gln Glu Thr Arg Arg Glu Arg
                860                 865                 870

Arg Arg Leu Met Ser Met Glu Met Asp
                875
```

<210> 137
<211> 434
<212> PRT
<213> Homo sapiens

<400> 137
Met Ser Arg Ser Arg His Leu Gly Lys Ile Arg Lys Arg Leu Glu
1 5 10 15

Asp Val Lys Ser Gln Trp Val Arg Pro Ala Arg Ala Asp Phe Ser
20 25 30

Asp Asn Glu Ser Ala Arg Leu Ala Thr Asp Ala Leu Leu Asp Gly
35 40 45

Gly Ser Glu Ala Tyr Trp Arg Val Leu Ser Gln Glu Gly Glu Val
50 55 60

Asp Phe Leu Ser Ser Val Glu Ala Gln Tyr Ile Gln Ala Gln Ala
65 70 75

Arg Glu Pro Pro Cys Pro Pro Asp Thr Leu Gly Gly Ala Glu Ala
80 85 90

Gly Pro Lys Gly Leu Asp Ser Ser Ser Leu Gln Ser Gly Thr Tyr
95 100 105

Phe Pro Val Ala Ser Glu Gly Ser Glu Pro Ala Leu Leu His Ser
110 115 120

Trp Ala Ser Ala Glu Lys Pro Tyr Leu Lys Glu Lys Ser Ser Ala
125 130 135

Thr Val Tyr Phe Gln Thr Val Lys His Asn Asn Ile Arg Asp Leu
140 145 150

Val Arg Arg Cys Ile Thr Arg Thr Ser Gln Val Leu Val Ile Leu
155 160 165

Met Asp Val Phe Thr Asp Val Glu Ile Phe Cys Asp Ile Leu Glu
170 175 180

Ala Ala Asn Lys Arg Gly Val Phe Val Cys Val Leu Leu Asp Gln
185 190 195

Gly Gly Val Lys Leu Phe Gln Glu Met Cys Asp Lys Val Gln Ile
200 205 210

Ser Asp Ser His Leu Lys Asn Ile Ser Ile Arg Ser Val Glu Gly
215 220 225

Glu Ile Tyr Cys Ala Lys Ser Gly Arg Lys Phe Ala Gly Gln Ile
230 235 240

Arg Glu Lys Phe Ile Ile Ser Asp Trp Arg Phe Val Leu Ser Gly
245 250 255

Ser Tyr Ser Phe Thr Trp Leu Cys Gly His Val His Arg Asn Ile
260 265 270

Leu Ser Lys Phe Thr Gly Gln Ala Val Glu Leu Phe Asp Glu Glu

```
                275                 280                 285

 Phe Arg His Leu Tyr Ala Ser Ser Lys Pro Val Met Gly Leu Lys
                290                 295                 300

 Ser Pro Arg Leu Val Ala Pro Val Pro Pro Gly Ala Ala Pro Ala
                305                 310                 315

 Asn Gly Arg Leu Ser Ser Ser Ser Gly Ser Ala Ser Asp Arg Thr
                320                 325                 330

 Ser Ser Asn Pro Phe Ser Gly Arg Ser Ala Gly Ser His Pro Gly
                335                 340                 345

 Thr Arg Ser Val Ser Ala Ser Ser Gly Pro Cys Ser Pro Ala Ala
                350                 355                 360

 Pro His Pro Pro Pro Pro Pro Arg Phe Gln Pro His Gln Gly Pro
                365                 370                 375

 Trp Gly Ala Pro Ser Pro Gln Ala His Leu Ser Pro Arg Pro His
                380                 385                 390

 Asp Gly Pro Pro Ala Ala Val Tyr Ser Asn Leu Gly Ala Tyr Arg
                395                 400                 405

 Pro Thr Arg Leu Gln Leu Glu Gln Leu Gly Leu Val Pro Arg Leu
                410                 415                 420

 Thr Pro Thr Trp Arg Pro Phe Leu Gln Ala Ser Pro His Phe
                425                 430

<210> 138
<211> 933
<212> PRT
<213> Homo sapiens

<400> 138
 Met Ala Leu Pro Gln Pro Pro Pro Pro Ala Gly Ala Arg Asp Pro
   1               5                  10                  15

 Val Thr Pro Glu His Trp Ala Ser Gly Pro Pro Ser Gln Ala Gln
                20                  25                  30

 Pro Leu Leu Arg Gln Glu Ala Lys Glu Glu Glu Glu Gly Glu Glu
                35                  40                  45

 Thr Gly Val Gln Gly Ala Trp Gly Thr Gly Thr Ala Glu Gln Arg
                50                  55                  60

 Arg Arg Gly Trp Gly Glu Ala Ala Glu Ser Ala Ala Ala Glu Glu
                65                  70                  75

 Gly Gln Ala Glu Val Gly Gly Ala Ala Ala Ala Gly Ser Gly Ser
                80                  85                  90

 Pro Ala Gly Gly Ala Gly Gly Gly Leu Gly Ser Trp Arg Pro Leu
                95                 100                 105

 Leu Ala Trp Leu Gln Arg Arg Gln Pro Gln Cys Cys Pro Cys Ala
                110                 115                 120
```

```
Ala Pro Leu Ser Arg Ser Ala Ala His Cys Cys His Gly Gly Thr
                125                 130                 135

Lys Met Ala Ala Leu Ala Tyr Asn Leu Gly Lys Arg Glu Ile Asn
                140                 145                 150

His Tyr Phe Ser Val Arg Ser Ala Lys Val Leu Ala Leu Val Ala
                155                 160                 165

Val Leu Leu Leu Ala Ala Cys His Leu Ala Ser Arg Arg Tyr Arg
                170                 175                 180

Gly Asn Asp Ser Cys Glu Tyr Leu Leu Ser Ser Gly Arg Phe Leu
                185                 190                 195

Gly Glu Lys Val Trp Gln Pro His Ser Cys Met Met His Lys Tyr
                200                 205                 210

Lys Ile Ser Glu Ala Lys Asn Cys Leu Val Asp Lys His Ile Ala
                215                 220                 225

Phe Ile Gly Asp Ser Arg Ile Arg Gln Leu Phe Tyr Ser Phe Val
                230                 235                 240

Lys Ile Ile Asn Pro Gln Phe Lys Glu Glu Gly Asn Lys His Glu
                245                 250                 255

Asn Ile Pro Phe Glu Asp Lys Thr Ala Ser Val Lys Val Asp Phe
                260                 265                 270

Leu Trp His Pro Glu Val Asn Gly Ser Met Lys Gln Cys Ile Lys
                275                 280                 285

Val Trp Thr Glu Asp Ser Ile Ala Lys Pro His Val Ile Val Ala
                290                 295                 300

Gly Ala Ala Thr Trp Ser Ile Lys Ile His Asn Gly Ser Ser Glu
                305                 310                 315

Ala Leu Ser Gln Tyr Lys Met Asn Ile Thr Ser Ile Ala Pro Leu
                320                 325                 330

Leu Glu Lys Leu Ala Lys Thr Ser Asp Val Tyr Trp Val Leu Gln
                335                 340                 345

Asp Pro Val Tyr Glu Asp Leu Leu Ser Glu Asn Arg Lys Met Ile
                350                 355                 360

Thr Asn Glu Lys Ile Asp Ala Tyr Asn Glu Ala Ala Val Ser Ile
                365                 370                 375

Leu Asn Ser Ser Thr Arg Asn Ser Lys Ser Asn Val Lys Met Phe
                380                 385                 390

Ser Val Ser Lys Leu Ile Ala Gln Glu Thr Ile Met Glu Ser Leu
                395                 400                 405

Asp Gly Leu His Leu Pro Glu Ser Ser Arg Glu Thr Thr Ala Met
                410                 415                 420
```

```
Ile Leu Met Asn Val Tyr Cys Asn Lys Ile Leu Lys Pro Val Asp
                425                 430                 435

Gly Ser Cys Cys Gln Pro Arg Pro Pro Val Thr Leu Ile Gln Lys
                440                 445                 450

Leu Ala Ala Cys Phe Phe Thr Leu Ser Ile Ile Gly Tyr Leu Ile
                455                 460                 465

Phe Tyr Ile Ile His Arg Asn Ala His Arg Lys Asn Lys Pro Cys
                470                 475                 480

Thr Asp Leu Glu Ser Gly Glu Glu Lys Lys Asn Ile Ile Asn Thr
                485                 490                 495

Pro Val Ser Ser Leu Glu Ile Leu Leu Gln Ser Phe Cys Lys Leu
                500                 505                 510

Gly Leu Ile Met Ala Tyr Phe Tyr Met Cys Asp Arg Ala Asn Leu
                515                 520                 525

Phe Met Lys Glu Asn Lys Phe Tyr Thr His Ser Ser Phe Phe Ile
                530                 535                 540

Pro Ile Ile Tyr Ile Leu Val Leu Gly Val Phe Tyr Asn Glu Asn
                545                 550                 555

Thr Lys Glu Thr Lys Val Leu Asn Arg Glu Gln Thr Asp Glu Trp
                560                 565                 570

Lys Gly Trp Met Gln Leu Val Ile Leu Ile Tyr His Ile Ser Gly
                575                 580                 585

Ala Ser Thr Phe Leu Pro Val Tyr Met His Ile Arg Val Leu Val
                590                 595                 600

Ala Ala Tyr Leu Phe Gln Thr Gly Tyr Gly His Phe Ser Tyr Phe
                605                 610                 615

Trp Ile Lys Gly Asp Phe Gly Ile Tyr Arg Val Cys Gln Val Leu
                620                 625                 630

Phe Arg Leu Asn Phe Leu Val Val Val Leu Cys Ile Val Met Asp
                635                 640                 645

Arg Pro Tyr Gln Phe Tyr Tyr Phe Val Pro Leu Val Thr Val Trp
                650                 655                 660

Phe Met Val Ile Tyr Val Thr Leu Ala Leu Trp Pro Gln Ile Ile
                665                 670                 675

Gln Lys Lys Ala Asn Gly Asn Cys Phe Trp His Phe Gly Leu Leu
                680                 685                 690

Leu Lys Leu Gly Phe Leu Leu Leu Phe Ile Cys Phe Leu Ala Tyr
                695                 700                 705

Ser Gln Gly Ala Phe Glu Lys Ile Phe Ser Leu Trp Pro Leu Ser
                710                 715                 720

Lys Cys Phe Glu Leu Lys Gly Asn Val Tyr Glu Trp Trp Phe Arg
```

```
                725                 730                 735

 Trp Arg Leu Asp Arg Tyr Val Val Phe His Gly Met Leu Phe Ala
                740                 745                 750

 Phe Ile Tyr Leu Ala Leu Gln Lys Arg Gln Ile Leu Ser Glu Gly
                755                 760                 765

 Lys Gly Glu Pro Leu Phe Ser Asn Lys Ile Ser Asn Phe Leu Leu
                770                 775                 780

 Phe Ile Ser Val Val Ser Phe Leu Thr Tyr Ser Ile Trp Ala Ser
                785                 790                 795

 Ser Cys Lys Asn Lys Ala Glu Cys Asn Glu Leu His Pro Ser Val
                800                 805                 810

 Ser Val Val Gln Ile Leu Ala Phe Ile Leu Ile Arg Asn Ile Pro
                815                 820                 825

 Gly Tyr Ala Arg Ser Val Tyr Ser Ser Phe Phe Ala Trp Phe Gly
                830                 835                 840

 Lys Ile Ser Leu Glu Leu Phe Ile Cys Gln Tyr His Ile Trp Leu
                845                 850                 855

 Ala Ala Asp Thr Arg Gly Ile Leu Val Leu Ile Pro Gly Asn Pro
                860                 865                 870

 Met Leu Asn Ile Ile Val Ser Thr Phe Ile Phe Val Cys Val Ala
                875                 880                 885

 His Glu Ile Ser Gln Ile Thr Asn Asp Leu Ala Gln Ile Ile Ile
                890                 895                 900

 Pro Lys Asp Asn Ser Ser Leu Leu Lys Arg Leu Ala Cys Ile Ala
                905                 910                 915

 Ala Phe Phe Cys Gly Leu Leu Ile Leu Ser Ser Ile Gln Asp Lys
                920                 925                 930

 Ser Lys His


<210> 139
<211> 320
<212> PRT
<213> Homo sapiens

<400> 139
 Met Tyr His Cys His Ser Gly Ser Lys Pro Thr Glu Lys Gly Ala
   1               5                  10                  15

 Asn Glu Tyr Ala Tyr Ala Lys Trp Lys Leu Cys Ser Ala Ser Ala
                20                  25                  30

 Ile Cys Phe Ile Phe Met Ile Ala Glu Val Val Gly Gly His Ile
                35                  40                  45

 Ala Gly Ser Leu Ala Val Val Thr Asp Ala Ala His Leu Leu Ile
                50                  55                  60
```

```
 Asp Leu Thr Ser Phe Leu Leu Ser Leu Phe Ser Leu Trp Leu Ser
                  65                  70                  75

 Ser Lys Pro Pro Ser Lys Arg Leu Thr Phe Gly Trp His Arg Ala
                  80                  85                  90

 Glu Ile Leu Gly Ala Leu Leu Ser Ile Leu Cys Ile Trp Val Val
                  95                 100                 105

 Thr Gly Val Leu Val Tyr Leu Ala Cys Glu Arg Leu Leu Tyr Pro
                 110                 115                 120

 Asp Tyr Gln Ile Gln Ala Thr Val Met Ile Ile Val Ser Ser Cys
                 125                 130                 135

 Ala Val Ala Ala Asn Ile Val Leu Thr Val Val Leu His Gln Arg
                 140                 145                 150

 Cys Leu Gly His Asn His Lys Glu Val Gln Ala Asn Ala Ser Val
                 155                 160                 165

 Arg Ala Ala Phe Val His Ala Leu Gly Asp Leu Phe Gln Ser Ile
                 170                 175                 180

 Ser Val Leu Ile Ser Ala Leu Ile Ile Tyr Phe Lys Pro Glu Tyr
                 185                 190                 195

 Lys Ile Ala Asp Pro Ile Cys Thr Phe Ile Phe Ser Ile Leu Val
                 200                 205                 210

 Leu Ala Ser Thr Ile Thr Ile Leu Lys Asp Phe Ser Ile Leu Leu
                 215                 220                 225

 Met Glu Gly Val Pro Lys Ser Leu Asn Tyr Ser Gly Val Lys Glu
                 230                 235                 240

 Leu Ile Leu Ala Val Asp Gly Val Leu Ser Val His Ser Leu His
                 245                 250                 255

 Ile Trp Ser Leu Thr Met Asn Gln Val Ile Leu Ser Ala His Val
                 260                 265                 270

 Ala Thr Ala Ala Ser Arg Asp Ser Gln Val Val Arg Arg Glu Ile
                 275                 280                 285

 Ala Lys Ala Leu Ser Lys Ser Phe Thr Met His Ser Leu Thr Ile
                 290                 295                 300

 Gln Met Glu Ser Pro Val Asp Gln Asp Pro Asp Cys Leu Phe Cys
                 305                 310                 315

 Glu Asp Pro Cys Asp
                 320

<210> 140
<211> 426
<212> PRT
<213> Homo sapiens

<400> 140
```

```
Met Val Leu Ser Val Pro Val Ile Ala Leu Gly Ala Thr Leu Gly
  1               5                  10                  15

Thr Ala Thr Ser Ile Leu Ala Leu Cys Gly Val Thr Cys Leu Cys
                20                  25                  30

Arg His Met His Pro Lys Lys Gly Leu Leu Pro Arg Asp Gln Asp
                35                  40                  45

Pro Asp Leu Glu Lys Ala Lys Pro Ser Leu Leu Gly Ser Ala Gln
                50                  55                  60

Gln Phe Asn Val Lys Lys Ser Thr Glu Pro Val Gln Pro Arg Ala
                65                  70                  75

Leu Leu Lys Phe Pro Asp Ile Tyr Gly Pro Arg Pro Ala Val Thr
                80                  85                  90

Ala Pro Glu Val Ile Asn Tyr Ala Asp Tyr Ser Leu Arg Ser Thr
                95                 100                 105

Glu Glu Pro Thr Ala Pro Ala Ser Pro Gln Pro Pro Asn Asp Ser
               110                 115                 120

Arg Leu Lys Arg Gln Val Thr Glu Glu Leu Phe Ile Leu Pro Gln
               125                 130                 135

Asn Gly Val Val Glu Asp Val Cys Val Met Glu Thr Trp Asn Pro
               140                 145                 150

Glu Lys Ala Ala Ser Trp Asn Gln Ala Pro Lys Leu His Tyr Cys
               155                 160                 165

Leu Asp Tyr Asp Cys Gln Lys Ala Glu Leu Phe Val Thr Arg Leu
               170                 175                 180

Glu Ala Val Thr Ser Asn His Asp Gly Gly Cys Asp Cys Tyr Val
               185                 190                 195

Gln Gly Ser Val Ala Asn Arg Thr Gly Ser Val Glu Ala Gln Thr
               200                 205                 210

Ala Leu Lys Lys Arg Gln Leu His Thr Thr Trp Glu Glu Gly Leu
               215                 220                 225

Val Leu Pro Leu Ala Glu Glu Glu Leu Pro Thr Ala Thr Leu Thr
               230                 235                 240

Leu Thr Leu Arg Thr Cys Asp Arg Phe Ser Arg His Ser Val Ala
               245                 250                 255

Gly Glu Leu Arg Leu Gly Leu Asp Gly Thr Ser Val Pro Leu Gly
               260                 265                 270

Ala Ala Gln Trp Gly Glu Leu Lys Thr Ser Ala Lys Glu Pro Ser
               275                 280                 285

Ala Gly Ala Gly Glu Val Leu Leu Ser Ile Ser Tyr Leu Pro Ala
               290                 295                 300

Ala Asn Arg Leu Leu Val Val Leu Ile Lys Ala Lys Asn Leu His
```

```
                305                 310                 315

 Ser Asn Gln Ser Lys Glu Leu Leu Gly Lys Asp Val Ser Val Lys
                320                 325                 330

 Val Thr Leu Lys His Gln Ala Arg Lys Leu Lys Lys Lys Gln Thr
                335                 340                 345

 Lys Arg Ala Lys His Lys Ile Asn Pro Val Trp Asn Glu Met Ile
                350                 355                 360

 Met Phe Glu Leu Pro Asp Asp Leu Leu Gln Ala Ser Ser Val Glu
                365                 370                 375

 Leu Glu Val Leu Gly Gln Asp Asp Ser Gly Gln Ser Cys Ala Leu
                380                 385                 390

 Gly His Cys Ser Leu Gly Leu His Thr Ser Gly Ser Glu Arg Ser
                395                 400                 405

 His Trp Glu Glu Met Leu Lys Asn Pro Arg Arg Gln Ile Ala Met
                410                 415                 420

 Trp His Gln Leu His Leu
                425

<210> 141
<211> 883
<212> PRT
<213> Homo sapiens

<400> 141
 Met Gln Lys Ile Met His Ile Ser Val Leu Leu Ser Pro Val Leu
   1               5                  10                  15

 Trp Gly Leu Ile Phe Gly Val Ser Ser Asn Ser Ile Gln Ile Gly
                20                  25                  30

 Gly Leu Phe Pro Arg Gly Ala Asp Gln Glu Tyr Ser Ala Phe Arg
                35                  40                  45

 Val Gly Met Val Gln Phe Ser Thr Ser Glu Phe Arg Leu Thr Pro
                50                  55                  60

 His Ile Asp Asn Leu Glu Val Ala Asn Ser Phe Ala Val Thr Asn
                65                  70                  75

 Ala Phe Cys Ser Gln Phe Ser Arg Gly Val Tyr Ala Ile Phe Gly
                80                  85                  90

 Phe Tyr Asp Lys Lys Ser Val Asn Thr Ile Thr Ser Phe Cys Gly
                95                 100                 105

 Thr Leu His Val Ser Phe Ile Thr Pro Ser Phe Pro Thr Asp Gly
               110                 115                 120

 Thr His Pro Phe Val Ile Gln Met Arg Pro Asp Leu Lys Gly Ala
               125                 130                 135

 Leu Leu Ser Leu Ile Glu Tyr Tyr Gln Trp Asp Lys Phe Ala Tyr
               140                 145                 150
```

```
Leu Tyr Asp Ser Asp Arg Gly Leu Ser Thr Leu Gln Ala Val Leu
                155                 160                 165

Asp Ser Ala Ala Glu Lys Lys Trp Gln Val Thr Ala Ile Asn Val
                170                 175                 180

Gly Asn Ile Asn Asn Asp Lys Lys Asp Glu Met Tyr Arg Ser Leu
                185                 190                 195

Phe Gln Asp Leu Glu Leu Lys Lys Glu Arg Arg Val Ile Leu Asp
                200                 205                 210

Cys Glu Arg Asp Lys Val Asn Asp Ile Val Asp Gln Val Ile Thr
                215                 220                 225

Ile Gly Lys His Val Lys Gly Tyr His Tyr Ile Ile Ala Asn Leu
                230                 235                 240

Glu Phe Thr Asp Gly Asp Leu Leu Lys Ile Gln Phe Gly Gly Ala
                245                 250                 255

Asn Val Ser Gly Phe Gln Ile Val Asp Tyr Asp Asp Ser Leu Val
                260                 265                 270

Ser Lys Phe Ile Glu Arg Trp Ser Thr Leu Glu Glu Lys Glu Tyr
                275                 280                 285

Pro Gly Ala His Thr Thr Thr Ile Lys Tyr Thr Ser Ala Leu Thr
                290                 295                 300

Tyr Asp Ala Val Gln Val Met Thr Glu Ala Phe Arg Asn Leu Arg
                305                 310                 315

Lys Gln Arg Ile Glu Ile Ser Arg Arg Gly Asn Ala Gly Asp Cys
                320                 325                 330

Leu Ala Asn Pro Ala Val Pro Trp Gly Gln Gly Val Glu Ile Glu
                335                 340                 345

Arg Ala Leu Lys Gln Val Gln Val Glu Gly Leu Ser Gly Asn Ile
                350                 355                 360

Lys Phe Asp Gln Asn Gly Lys Arg Ile Asn Tyr Thr Ile Asn Ile
                365                 370                 375

Met Glu Leu Lys Thr Asn Gly Pro Arg Lys Ile Gly Tyr Trp Ser
                380                 385                 390

Glu Val Asp Lys Met Val Val Thr Leu Thr Glu Leu Pro Ser Gly
                395                 400                 405

Asn Asp Thr Ser Gly Leu Glu Asn Lys Thr Val Val Val Thr Thr
                410                 415                 420

Ile Leu Glu Ser Pro Tyr Val Met Met Lys Lys Asn His Glu Met
                425                 430                 435

Leu Glu Gly Asn Glu Arg Tyr Glu Gly Tyr Cys Val Asp Leu Ala
                440                 445                 450
```

```
Ala Glu Ile Ala Lys His Cys Gly Phe Lys Tyr Lys Leu Thr Ile
                455                 460                 465

Val Gly Asp Gly Lys Tyr Gly Ala Arg Asp Ala Asp Thr Lys Ile
                470                 475                 480

Trp Asn Gly Met Val Gly Glu Leu Val Tyr Gly Lys Ala Asp Ile
                485                 490                 495

Ala Ile Ala Pro Leu Thr Ile Thr Leu Val Arg Glu Glu Val Ile
                500                 505                 510

Asp Phe Ser Lys Pro Phe Met Ser Leu Gly Ile Ser Ile Met Ile
                515                 520                 525

Lys Lys Pro Gln Lys Ser Lys Pro Gly Val Phe Ser Phe Leu Asp
                530                 535                 540

Pro Leu Ala Tyr Glu Ile Trp Met Cys Ile Val Phe Ala Tyr Ile
                545                 550                 555

Gly Val Ser Val Val Leu Phe Leu Val Ser Arg Phe Ser Pro Tyr
                560                 565                 570

Glu Trp His Thr Glu Glu Phe Glu Asp Gly Arg Glu Thr Gln Ser
                575                 580                 585

Ser Glu Ser Thr Asn Glu Phe Gly Ile Phe Asn Ser Leu Trp Phe
                590                 595                 600

Ser Leu Gly Ala Phe Met Arg Gln Gly Cys Asp Ile Ser Pro Arg
                605                 610                 615

Ser Leu Ser Gly Arg Ile Val Gly Gly Val Trp Trp Phe Phe Thr
                620                 625                 630

Leu Ile Ile Ile Ser Ser Tyr Thr Ala Asn Leu Ala Ala Phe Leu
                635                 640                 645

Thr Val Glu Arg Met Val Ser Pro Ile Glu Ser Ala Glu Asp Leu
                650                 655                 660

Ser Lys Gln Thr Glu Ile Ala Tyr Gly Thr Leu Asp Ser Gly Ser
                665                 670                 675

Thr Lys Glu Phe Phe Arg Arg Ser Lys Ile Ala Val Phe Asp Lys
                680                 685                 690

Met Trp Thr Tyr Met Arg Ser Ala Glu Pro Ser Val Phe Val Arg
                695                 700                 705

Thr Thr Ala Glu Gly Val Ala Arg Val Arg Lys Ser Lys Gly Lys
                710                 715                 720

Tyr Ala Tyr Leu Leu Glu Ser Thr Met Asn Glu Tyr Ile Glu Gln
                725                 730                 735

Arg Lys Pro Cys Asp Thr Met Lys Val Gly Gly Asn Leu Asp Ser
                740                 745                 750

Lys Gly Tyr Gly Ile Ala Thr Pro Lys Gly Ser Ser Leu Gly Thr
```

```
                 755                 760                 765

 Pro Val Asn Leu Ala Val Leu Lys Leu Ser Glu Gln Gly Val Leu
                 770                 775                 780

 Asp Lys Leu Lys Asn Lys Trp Trp Tyr Asp Lys Gly Glu Cys Gly
                 785                 790                 795

 Ala Lys Asp Ser Gly Ser Lys Glu Lys Thr Ser Ala Leu Ser Leu
                 800                 805                 810

 Ser Asn Val Ala Gly Val Phe Tyr Ile Leu Val Gly Gly Leu Gly
                 815                 820                 825

 Leu Ala Met Leu Val Ala Leu Ile Glu Phe Cys Tyr Lys Ser Arg
                 830                 835                 840

 Ala Glu Ala Lys Arg Met Lys Val Ala Lys Asn Ala Gln Asn Ile
                 845                 850                 855

 Asn Pro Ser Ser Ser Gln Asn Ser Gln Asn Phe Ala Thr Tyr Lys
                 860                 865                 870

 Glu Gly Tyr Asn Val Tyr Gly Ile Glu Ser Val Lys Ile
                 875                 880

<210> 142
<211> 357
<212> PRT
<213> Homo sapiens

<400> 142
 Met Ala Thr Thr Val Pro Asp Gly Cys Arg Asn Gly Leu Lys Ser
   1               5                  10                  15

 Lys Tyr Tyr Arg Leu Cys Asp Lys Ala Glu Ala Trp Gly Ile Val
                  20                  25                  30

 Leu Glu Thr Val Ala Thr Ala Gly Val Val Thr Ser Val Ala Phe
                  35                  40                  45

 Met Leu Thr Leu Pro Ile Leu Val Cys Lys Val Gln Asp Ser Asn
                  50                  55                  60

 Arg Arg Lys Met Leu Pro Thr Gln Phe Leu Phe Leu Leu Gly Val
                  65                  70                  75

 Leu Gly Ile Phe Gly Leu Thr Phe Ala Phe Ile Ile Gly Leu Asp
                  80                  85                  90

 Gly Ser Thr Gly Pro Thr Arg Phe Phe Leu Phe Gly Ile Leu Phe
                  95                 100                 105

 Ser Ile Cys Phe Ser Cys Leu Leu Ala His Ala Val Ser Leu Thr
                 110                 115                 120

 Lys Leu Val Arg Gly Arg Lys Pro Leu Ser Leu Leu Val Ile Leu
                 125                 130                 135

 Gly Leu Ala Val Gly Phe Ser Leu Val Gln Asp Val Ile Ala Ile
                 140                 145                 150
```

```
 Glu Tyr Ile Val Leu Thr Met Asn Arg Thr Asn Val Asn Val Phe
                 155                 160                 165

 Ser Glu Leu Ser Ala Pro Arg Arg Asn Glu Asp Phe Val Leu Leu
                 170                 175                 180

 Leu Thr Tyr Val Leu Phe Leu Met Ala Leu Thr Phe Leu Met Ser
                 185                 190                 195

 Ser Phe Thr Phe Cys Gly Ser Phe Thr Gly Trp Lys Arg His Gly
                 200                 205                 210

 Ala His Ile Tyr Leu Thr Met Leu Leu Ser Ile Ala Ile Trp Val
                 215                 220                 225

 Ala Trp Ile Thr Leu Leu Met Leu Pro Asp Phe Asp Arg Arg Trp
                 230                 235                 240

 Asp Asp Thr Ile Leu Ser Ser Ala Leu Ala Ala Asn Gly Trp Val
                 245                 250                 255

 Phe Leu Leu Ala Tyr Val Ser Pro Glu Phe Trp Leu Leu Thr Lys
                 260                 265                 270

 Gln Arg Asn Pro Met Asp Tyr Pro Val Glu Asp Ala Phe Cys Lys
                 275                 280                 285

 Pro Gln Leu Val Lys Lys Ser Tyr Gly Val Glu Asn Arg Ala Tyr
                 290                 295                 300

 Ser Gln Glu Glu Ile Thr Gln Gly Phe Glu Glu Thr Gly Asp Thr
                 305                 310                 315

 Leu Tyr Ala Pro Tyr Ser Thr His Phe Gln Leu Gln Asn Gln Pro
                 320                 325                 330

 Pro Gln Lys Glu Phe Ser Ile Pro Arg Ala His Ala Trp Pro Ser
                 335                 340                 345

 Pro Tyr Lys Asp Tyr Glu Val Lys Lys Glu Gly Ser
                 350                 355

<210> 143
<211> 222
<212> PRT
<213> Homo sapiens

<400> 143
 Met Leu Pro Ala Gln Glu Ala Ala Lys Leu Tyr His Thr Asn Tyr
   1               5                  10                  15

 Val Arg Asn Ser Arg Ala Ile Gly Val Leu Trp Ala Ile Phe Thr
                  20                  25                  30

 Ile Cys Phe Ala Ile Val Asn Val Val Cys Phe Ile Gln Pro Tyr
                  35                  40                  45

 Trp Ile Gly Asp Gly Val Asp Thr Pro Gln Ala Gly Tyr Phe Gly
                  50                  55                  60
```

```
 Leu Phe His Tyr Cys Ile Gly Asn Gly Phe Ser Arg Glu Leu Thr
                  65                  70                  75

 Cys Arg Gly Ser Phe Thr Asp Phe Ser Thr Leu Pro Ser Gly Ala
                  80                  85                  90

 Phe Lys Ala Ala Ser Phe Phe Ile Gly Leu Ser Met Met Leu Ile
                  95                 100                 105

 Ile Ala Cys Ile Ile Cys Phe Thr Leu Phe Phe Phe Cys Asn Thr
                 110                 115                 120

 Ala Thr Val Tyr Lys Ile Cys Ala Trp Met Gln Leu Thr Ser Ala
                 125                 130                 135

 Ala Cys Leu Val Leu Gly Cys Met Ile Phe Pro Asp Gly Trp Asp
                 140                 145                 150

 Ser Asp Glu Val Lys Arg Met Cys Gly Glu Lys Thr Asp Lys Tyr
                 155                 160                 165

 Thr Leu Gly Ala Cys Ser Val Arg Trp Ala Tyr Ile Leu Ala Ile
                 170                 175                 180

 Ile Gly Ile Leu Asp Ala Leu Ile Leu Ser Phe Leu Ala Phe Val
                 185                 190                 195

 Leu Gly Asn Arg Gln Asp Ser Leu Met Ala Glu Glu Leu Lys Ala
                 200                 205                 210

 Glu Asn Lys Val Leu Leu Ser Gln Tyr Ser Leu Glu
                 215                 220

<210> 144
<211> 236
<212> PRT
<213> Homo sapiens

<400> 144
 Met Pro Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Met
   1               5                  10                  15

 Leu Pro Ala Gln Glu Ala Ala Lys Leu Tyr His Thr Asn Tyr Val
                  20                  25                  30

 Arg Asn Ser Arg Ala Ile Gly Val Leu Trp Ala Ile Phe Thr Ile
                  35                  40                  45

 Cys Phe Ala Ile Val Asn Val Val Cys Phe Ile Gln Pro Tyr Trp
                  50                  55                  60

 Ile Gly Asp Gly Val Asp Thr Pro Gln Ala Gly Tyr Phe Gly Leu
                  65                  70                  75

 Phe His Tyr Cys Ile Gly Asn Gly Phe Ser Arg Glu Leu Thr Cys
                  80                  85                  90

 Arg Gly Ser Phe Thr Asp Phe Ser Thr Leu Pro Ser Gly Ala Phe
                  95                 100                 105

 Lys Ala Ala Ser Phe Phe Ile Gly Leu Ser Met Met Leu Ile Ile
```

```
                110                 115                 120

 Ala Cys Ile Ile Cys Phe Thr Leu Phe Phe Phe Cys Asn Thr Ala
                125                 130                 135

 Thr Val Tyr Lys Ile Cys Ala Trp Met Gln Leu Thr Ser Ala Ala
                140                 145                 150

 Cys Leu Val Leu Gly Cys Met Ile Phe Pro Asp Gly Trp Asp Ser
                155                 160                 165

 Asp Glu Val Lys Arg Met Cys Gly Glu Lys Thr Asp Lys Tyr Thr
                170                 175                 180

 Leu Gly Ala Cys Ser Val Arg Trp Ala Tyr Ile Leu Ala Ile Ile
                185                 190                 195

 Gly Ile Leu Asp Ala Leu Ile Leu Ser Phe Leu Ala Phe Val Leu
                200                 205                 210

 Gly Asn Arg Gln Asp Ser Leu Met Ala Glu Glu Leu Lys Ala Glu
                215                 220                 225

 Asn Lys Val Leu Leu Ser Gln Tyr Ser Leu Glu
                230                 235

<210> 145
<211> 236
<212> PRT
<213> Homo sapiens

<400> 145
 Met Pro Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Met
   1               5                  10                  15

 Leu Pro Ala Gln Glu Ala Ala Lys Leu Tyr His Thr Asn Tyr Val
                 20                  25                  30

 Arg Asn Ser Arg Ala Ile Gly Val Leu Trp Ala Ile Phe Thr Ile
                 35                  40                  45

 Cys Phe Ala Ile Val Asn Val Val Cys Phe Ile Gln Pro Tyr Trp
                 50                  55                  60

 Ile Gly Asp Gly Val Asp Thr Pro Gln Ala Gly Tyr Phe Gly Leu
                 65                  70                  75

 Phe His Tyr Cys Ile Gly Asn Gly Phe Ser Arg Glu Leu Thr Cys
                 80                  85                  90

 Arg Gly Ser Phe Thr Asp Phe Ser Thr Leu Pro Ser Gly Ala Phe
                 95                 100                 105

 Lys Ala Ala Ser Phe Phe Ile Gly Leu Ser Met Met Leu Ile Ile
                110                 115                 120

 Ala Cys Ile Ile Cys Phe Thr Leu Phe Phe Phe Cys Asn Thr Ala
                125                 130                 135

 Thr Val Tyr Lys Ile Cys Ala Trp Met Gln Leu Thr Ser Ala Ala
                140                 145                 150
```

```
 Cys Leu Val Leu Gly Cys Met Ile Phe Pro Asp Gly Trp Asp Ser
                 155                 160                 165

 Asp Glu Val Lys Arg Met Cys Gly Glu Lys Thr Asp Lys Tyr Thr
                 170                 175                 180

 Leu Gly Ala Cys Ser Val Arg Trp Ala Tyr Ile Leu Ala Ile Ile
                 185                 190                 195

 Gly Ile Leu Asp Ala Leu Ile Leu Ser Phe Leu Ala Phe Val Leu
                 200                 205                 210

 Gly Asn Arg Gln Asp Ser Leu Met Ala Glu Glu Leu Lys Ala Glu
                 215                 220                 225

 Asn Lys Val Leu Leu Ser Gln Tyr Ser Leu Glu
                 230                 235

<210> 146
<211> 408
<212> PRT
<213> Homo sapiens

<400> 146
 Arg Pro Leu Arg Leu Pro Cys Arg Gly Pro Arg Ala Ala Gly Thr
   1               5                  10                  15

 Val Ser Ser Glu Ser Leu Gly His Pro Ala Pro Gly Gln Val Thr
                  20                  25                  30

 Asn Arg Ser Pro Gln Phe Pro Met Glu Leu Phe Gln Ser Ser Pro
                  35                  40                  45

 Arg Pro Pro Leu Gln Pro Val Cys Val Cys Pro Gly Leu Glu Leu
                  50                  55                  60

 Cys Asp Pro Pro His Arg Leu Ala Phe Ser Met Gly Ala Gly Leu
                  65                  70                  75

 Phe Ser Val Val Gly Thr Leu Leu Leu Pro Gly Leu Ala Ala Leu
                  80                  85                  90

 Val Gln Asp Trp Arg Leu Leu Gln Gly Leu Gly Ala Leu Met Ser
                  95                 100                 105

 Gly Leu Leu Leu Leu Phe Trp Gly Phe Pro Ala Leu Phe Pro Glu
                 110                 115                 120

 Ser Pro Cys Trp Leu Leu Ala Thr Gly Gln Val Ala Arg Ala Arg
                 125                 130                 135

 Lys Ile Leu Trp Arg Phe Ala Glu Ala Ser Gly Val Gly Pro Gly
                 140                 145                 150

 Asp Ser Ser Leu Glu Glu Asn Ser Leu Ala Thr Glu Leu Thr Met
                 155                 160                 165

 Leu Ser Ala Arg Ser Pro Gln Pro Arg Tyr His Ser Pro Leu Gly
                 170                 175                 180
```

```
 Leu Leu Arg Thr Arg Val Thr Trp Arg Asn Gly Leu Ile Leu Gly
                 185                 190                 195

 Phe Ser Ser Leu Val Gly Gly Gly Ile Arg Ala Ser Phe Arg Arg
                 200                 205                 210

 Ser Leu Ala Pro Gln Val Pro Thr Phe Tyr Leu Pro Tyr Phe Leu
                 215                 220                 225

 Glu Ala Gly Leu Glu Ala Ala Ala Leu Val Phe Leu Leu Leu Thr
                 230                 235                 240

 Ala Asp Cys Cys Gly Arg Arg Pro Val Leu Leu Leu Gly Thr Met
                 245                 250                 255

 Val Thr Gly Leu Ala Ser Leu Leu Leu Leu Ala Gly Ala Gln Tyr
                 260                 265                 270

 Leu Pro Gly Trp Thr Val Leu Phe Leu Ser Val Leu Gly Leu Leu
                 275                 280                 285

 Ala Ser Arg Ala Val Ser Ala Leu Ser Ser Leu Phe Ala Ala Glu
                 290                 295                 300

 Val Phe Pro Thr Val Ile Arg Gly Ala Gly Leu Gly Leu Val Leu
                 305                 310                 315

 Gly Ala Gly Phe Leu Gly Gln Ala Ala Gly Pro Leu Asp Thr Leu
                 320                 325                 330

 His Gly Arg Gln Gly Phe Phe Leu Gln Gln Val Val Phe Ala Ser
                 335                 340                 345

 Leu Ala Val Leu Ala Leu Leu Cys Val Leu Leu Leu Pro Glu Ser
                 350                 355                 360

 Arg Ser Arg Gly Leu Pro Gln Ser Leu Gln Asp Ala Asp Arg Leu
                 365                 370                 375

 Arg Arg Ser Pro Leu Leu Arg Gly Arg Pro Arg Gln Asp His Leu
                 380                 385                 390

 Pro Leu Leu Pro Pro Ser Asn Ser Tyr Trp Ala Gly His Thr Pro
                 395                 400                 405

 Glu Gln His


<210> 147
<211> 1152
<212> PRT
<213> Homo sapiens

<220>
<221> unsure
<222> 382
<223> unknown amino acid

<400> 147
 Met Pro Phe Glu Gln Cys His Leu Ser Trp Lys Asn Val Arg Arg
   1               5                  10                  15
```

```
Leu Tyr Asn Ala Val Ser Lys Glu Leu Val Gly Glu Phe Leu Gln
                 20                  25                  30

Phe Val Gln Leu Asp Lys Glu Ala Ser Asp Pro Phe Ser Leu Asn
                 35                  40                  45

Glu Leu Leu Asp Glu Leu Ser Arg Lys Gln Lys Glu Glu Leu Trp
                 50                  55                  60

Gln Arg Leu Lys Asn Leu Leu Thr Asp Val Leu Leu Glu Ser Pro
                 65                  70                  75

Val Asp Gly Trp Gln Val Val Glu Ala Gln Gly Glu Asp Asn Met
                 80                  85                  90

Glu Thr Glu His Gly Ser Lys Met Arg Lys Ser Ile Glu Ile Ile
                 95                 100                 105

Tyr Ala Ile Thr Ser Val Ile Leu Ala Ser Val Ser Val Ile Asn
                110                 115                 120

Glu Ser Glu Asn Tyr Glu Ala Leu Leu Glu Cys Val Ile Ile Leu
                125                 130                 135

Asn Gly Ile Leu Tyr Ala Leu Pro Glu Ser Glu Arg Lys Leu Gln
                140                 145                 150

Ser Ser Ile Gln Asp Leu Cys Val Thr Trp Trp Glu Lys Gly Leu
                155                 160                 165

Pro Ala Lys Glu Asp Thr Gly Lys Thr Ala Phe Val Met Leu Leu
                170                 175                 180

Arg Arg Ser Leu Glu Thr Lys Thr Gly Ala Asp Val Cys Arg Leu
                185                 190                 195

Trp Arg Ile His Gln Ala Leu Tyr Cys Phe Asp Tyr Asp Leu Glu
                200                 205                 210

Glu Ser Gly Glu Ile Lys Asp Met Leu Leu Glu Cys Phe Ile Asn
                215                 220                 225

Ile Asn Tyr Ile Lys Lys Glu Glu Gly Arg Arg Phe Leu Ser Cys
                230                 235                 240

Leu Phe Asn Trp Asn Ile Asn Phe Ile Lys Met Ile His Gly Thr
                245                 250                 255

Ile Lys Asn Gln Leu Gln Gly Leu Gln Lys Ser Leu Met Val Tyr
                260                 265                 270

Ile Ala Glu Ile Tyr Phe Arg Ala Trp Lys Lys Ala Ser Gly Lys
                275                 280                 285

Ile Leu Glu Ala Ile Glu Asn Asp Cys Ile Gln Asp Phe Met Phe
                290                 295                 300

His Gly Ile His Leu Pro Arg Arg Ser Pro Val His Ser Lys Val
                305                 310                 315
```

```
Arg Glu Val Leu Ser Tyr Phe His His Gln Lys Lys Val Arg Gln
                320                 325                 330

Gly Val Glu Glu Met Leu Tyr Arg Leu Tyr Lys Pro Ile Leu Trp
                335                 340                 345

Arg Gly Leu Lys Ala Arg Asn Ser Glu Val Arg Ser Asn Ala Ala
                350                 355                 360

Leu Leu Phe Val Glu Ala Phe Pro Ile Arg Asp Pro Asn Leu His
                365                 370                 375

Ala Ile Glu Met Asp Ser Xaa Ile Gln Lys Gln Phe Glu Glu Leu
                380                 385                 390

Tyr Ser Leu Leu Glu Asp Pro Tyr Pro Met Val Arg Ser Thr Gly
                395                 400                 405

Ile Leu Gly Val Cys Lys Ile Thr Ser Lys Tyr Trp Glu Met Met
                410                 415                 420

Pro Pro Thr Ile Leu Ile Asp Leu Leu Lys Lys Val Thr Gly Glu
                425                 430                 435

Leu Ala Phe Asp Thr Ser Ser Ala Asp Val Arg Cys Ser Val Phe
                440                 445                 450

Lys Cys Leu Pro Met Ile Leu Asp Asn Lys Leu Ser His Pro Leu
                455                 460                 465

Leu Glu Gln Leu Leu Pro Ala Leu Arg Tyr Ser Leu His Asp Asn
                470                 475                 480

Ser Glu Lys Val Arg Val Ala Phe Val Asp Met Leu Leu Lys Ile
                485                 490                 495

Lys Ala Val Arg Ala Ala Lys Phe Trp Lys Ile Cys Pro Met Glu
                500                 505                 510

His Ile Leu Val Arg Leu Glu Thr Asp Ser Arg Pro Val Ser Arg
                515                 520                 525

Arg Leu Val Ser Leu Ile Phe Asn Ser Phe Leu Pro Val Asn Gln
                530                 535                 540

Pro Glu Glu Val Trp Cys Glu Arg Cys Val Thr Leu Val Gln Met
                545                 550                 555

Asn His Ala Ala Ala Arg Arg Phe Tyr Gln Tyr Ala His Glu His
                560                 565                 570

Thr Ala Cys Thr Asn Ile Ala Lys Leu Ile His Val Ile Arg His
                575                 580                 585

Cys Leu Asn Ala Cys Ile Gln Arg Ala Val Arg Glu Pro Pro Glu
                590                 595                 600

Asp Glu Glu Glu Glu Asp Gly Arg Glu Lys Glu Asn Val Thr Val
                605                 610                 615

Leu Asp Lys Thr Leu Ser Val Asn Asp Val Ala Cys Met Ala Gly
```

```
                620                 625                 630

Leu Leu Glu Ile Ile Val Ile Leu Trp Lys Ser Ile Asp Arg Ser
                635                 640                 645

Met Glu Asn Asn Lys Glu Ala Lys Leu Tyr Thr Ile Asn Lys Phe
                650                 655                 660

Ala Ser Val Leu Pro Glu Tyr Leu Lys Val Phe Lys Asp Asp Arg
                665                 670                 675

Cys Lys Ile Pro Leu Phe Met Leu Met Ser Phe Met Pro Ala Ser
                680                 685                 690

Ala Val Pro Pro Phe Ser Cys Gly Val Ile Ser Thr Leu Arg Ser
                695                 700                 705

Arg Glu Glu Gly Ala Val Asp Lys Ser Tyr Cys Thr Leu Leu Asp
                710                 715                 720

Cys Leu Cys Ser Trp Gly Gln Val Gly His Ile Leu Glu Leu Val
                725                 730                 735

Asp Asn Trp Leu Pro Thr Glu His Ala Gln Ala Lys Ser Asn Thr
                740                 745                 750

Ala Ser Lys Gly Arg Val Gln Ile His Asp Thr Arg Pro Val Lys
                755                 760                 765

Pro Glu Leu Ala Leu Val Tyr Ile Glu Tyr Leu Leu Thr His Pro
                770                 775                 780

Lys Asn Arg Glu Cys Leu Leu Ser Ala Pro Arg Lys Lys Leu Asn
                785                 790                 795

His Leu Leu Lys Ala Leu Glu Thr Ser Lys Ala Asp Leu Glu Ser
                800                 805                 810

Leu Leu Gln Thr Pro Gly Gly Lys Pro Arg Gly Phe Ser Glu Ala
                815                 820                 825

Ala Ala Pro Arg Ala Phe Gly Leu His Cys Arg Leu Ser Ile His
                830                 835                 840

Leu Gln His Lys Phe Cys Ser Glu Gly Lys Val Tyr Leu Ser Met
                845                 850                 855

Leu Glu Asp Thr Gly Phe Trp Leu Glu Ser Lys Ile Leu Ser Phe
                860                 865                 870

Ile Gln Asp Gln Glu Glu Asp Tyr Leu Lys Leu His Arg Val Ile
                875                 880                 885

Tyr Gln Gln Ile Ile Gln Thr Tyr Leu Thr Val Cys Lys Asp Val
                890                 895                 900

Val Met Val Gly Leu Gly Asp His Gln Phe Gln Met Gln Leu Leu
                905                 910                 915

Gln Arg Ser Leu Gly Ile Met Gln Thr Val Lys Gly Phe Phe Tyr
                920                 925                 930
```

```
 Val Ser Leu Leu Leu Asp Ile Leu Lys Glu Ile Thr Gly Ser Ser
                 935                 940                 945

 Leu Ile Gln Lys Thr Asp Ser Asp Glu Glu Val Ala Met Leu Leu
                 950                 955                 960

 Asp Thr Val Gln Lys Val Phe Gln Lys Met Leu Glu Cys Ile Ala
                 965                 970                 975

 Arg Ser Phe Arg Lys Gln Pro Glu Glu Gly Leu Arg Leu Leu Tyr
                 980                 985                 990

 Ser Val Gln Arg Pro Leu His Glu Phe Ile Thr Ala Val Gln Ser
                 995                1000                1005

 Arg His Thr Asp Thr Pro Val His Arg Gly Val Leu Ser Thr Leu
                1010                1015                1020

 Ile Ala Gly Pro Val Val Glu Ile Ser His Gln Leu Arg Lys Val
                1025                1030                1035

 Ser Asp Val Glu Glu Leu Thr Pro Pro Glu His Leu Ser Asp Leu
                1040                1045                1050

 Pro Pro Phe Ser Arg Cys Leu Ile Gly Ile Ile Ile Lys Ser Ser
                1055                1060                1065

 Asn Val Val Arg Ser Phe Leu Asp Glu Leu Lys Ala Cys Val Ala
                1070                1075                1080

 Ser Asn Asp Ile Glu Gly Ile Val Cys Leu Thr Ala Ala Val His
                1085                1090                1095

 Ile Ile Leu Val Ile Asn Ala Gly Lys His Lys Ser Ser Lys Val
                1100                1105                1110

 Arg Glu Val Ala Ala Thr Val His Arg Lys Leu Lys Thr Phe Met
                1115                1120                1125

 Glu Ile Thr Leu Glu Glu Asp Ser Ile Glu Arg Phe Leu Tyr Glu
                1130                1135                1140

 Ser Ser Ser Arg Thr Leu Gly Glu Leu Leu Asn Ser
                1145                1150

<210> 148
<211> 1143
<212> PRT
<213> Homo sapiens

<400> 148
 Met Glu Lys Arg Glu Thr Phe Val Gln Ala Val Ser Lys Glu Leu
   1               5                  10                  15

 Val Gly Glu Phe Leu Gln Phe Val Gln Leu Asp Lys Glu Ala Ser
                  20                  25                  30

 Asp Pro Phe Ser Leu Asn Glu Leu Leu Asp Glu Leu Ser Arg Lys
                  35                  40                  45
```

```
Gln Lys Glu Glu Leu Trp Gln Arg Leu Lys Asn Leu Leu Thr Asp
                 50                  55                  60

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

Gln Gly Glu Asp Asn Met Glu Thr Glu His Gly Ser Lys Met Arg
                 80                  85                  90

Lys Ser Ile Glu Ile Ile Tyr Ala Ile Thr Ser Val Ile Leu Ala
                 95                 100                 105

Ser Val Ser Val Ile Asn Glu Ser Glu Asn Tyr Glu Ala Leu Leu
                110                 115                 120

Glu Cys Val Ile Ile Leu Asn Gly Ile Leu Tyr Ala Leu Pro Glu
                125                 130                 135

Ser Glu Arg Lys Leu Gln Ser Ser Ile Gln Asp Leu Cys Val Thr
                140                 145                 150

Trp Trp Glu Lys Gly Leu Pro Ala Lys Glu Asp Thr Gly Lys Thr
                155                 160                 165

Ala Phe Val Met Leu Leu Arg Arg Ser Leu Glu Thr Lys Thr Gly
                170                 175                 180

Ala Asp Val Cys Arg Leu Trp Arg Ile His Gln Ala Leu Tyr Cys
                185                 190                 195

Phe Asp Tyr Asp Leu Glu Glu Ser Gly Glu Ile Lys Asp Met Leu
                200                 205                 210

Leu Glu Cys Phe Ile Asn Ile Asn Tyr Ile Lys Lys Glu Glu Gly
                215                 220                 225

Arg Arg Phe Leu Ser Cys Leu Phe Asn Trp Asn Ile Asn Phe Ile
                230                 235                 240

Lys Met Ile His Gly Thr Ile Lys Asn Gln Leu Gln Gly Leu Gln
                245                 250                 255

Lys Ser Leu Met Val Tyr Ile Ala Glu Ile Tyr Phe Arg Ala Trp
                260                 265                 270

Lys Lys Ala Ser Gly Lys Ile Leu Glu Ala Ile Glu Asn Asp Cys
                275                 280                 285

Ile Gln Asp Phe Met Phe His Gly Ile His Leu Pro Arg Arg Ser
                290                 295                 300

Pro Val His Ser Lys Val Arg Glu Val Leu Ser Tyr Phe His His
                305                 310                 315

Gln Lys Lys Val Arg Gln Gly Val Glu Glu Met Leu Tyr Arg Leu
                320                 325                 330

Tyr Lys Pro Ile Leu Trp Arg Gly Leu Lys Ala Arg Asn Ser Glu
                335                 340                 345

Val Arg Ser Asn Ala Ala Leu Leu Phe Val Glu Ala Phe Pro Ile
```

```
                350                 355                 360

Arg Asp Pro Asn Leu His Ala Ile Glu Met Asp Ser Glu Ile Gln
                365                 370                 375

Lys Gln Phe Glu Glu Leu Tyr Ser Leu Leu Glu Asp Pro Tyr Pro
                380                 385                 390

Met Val Arg Ser Thr Gly Ile Leu Gly Val Cys Lys Ile Thr Ser
                395                 400                 405

Lys Tyr Trp Glu Met Met Pro Pro Thr Ile Leu Ile Asp Leu Leu
                410                 415                 420

Lys Lys Val Thr Gly Glu Leu Ala Phe Asp Thr Ser Ser Ala Asp
                425                 430                 435

Val Arg Cys Ser Val Phe Lys Cys Leu Pro Met Ile Leu Asp Asn
                440                 445                 450

Lys Leu Ser His Pro Leu Leu Glu Gln Leu Leu Pro Ala Leu Arg
                455                 460                 465

Tyr Ser Leu His Asp Asn Ser Glu Lys Val Arg Val Ala Phe Val
                470                 475                 480

Asp Met Leu Leu Lys Ile Lys Ala Val Arg Ala Ala Lys Phe Trp
                485                 490                 495

Lys Ile Cys Pro Met Glu His Ile Leu Val Arg Leu Glu Thr Asp
                500                 505                 510

Ser Arg Pro Val Ser Arg Arg Leu Val Ser Leu Ile Phe Asn Ser
                515                 520                 525

Phe Leu Pro Val Asn Gln Pro Glu Glu Val Trp Cys Glu Arg Cys
                530                 535                 540

Val Thr Leu Val Gln Met Asn His Ala Ala Ala Arg Arg Phe Tyr
                545                 550                 555

Gln Tyr Ala His Glu His Thr Ala Cys Thr Asn Ile Ala Lys Leu
                560                 565                 570

Ile His Val Ile Arg His Cys Leu Asn Ala Cys Ile Gln Arg Ala
                575                 580                 585

Val Arg Glu Pro Pro Glu Asp Glu Glu Glu Glu Asp Gly Arg Glu
                590                 595                 600

Lys Glu Asn Val Thr Val Leu Asp Lys Thr Leu Ser Val Asn Asp
                605                 610                 615

Val Ala Cys Met Ala Gly Leu Leu Glu Ile Ile Val Ile Leu Trp
                620                 625                 630

Lys Ser Ile Asp Arg Ser Met Glu Asn Asn Lys Glu Ala Lys Leu
                635                 640                 645

Tyr Thr Ile Asn Lys Phe Ala Ser Val Leu Pro Glu Tyr Leu Lys
                650                 655                 660
```

```
Val Phe Lys Asp Asp Arg Cys Lys Ile Pro Leu Phe Met Leu Met
                665                 670                 675

Ser Phe Met Pro Ala Ser Ala Val Pro Pro Phe Ser Cys Gly Val
                680                 685                 690

Ile Ser Thr Leu Arg Ser Arg Glu Glu Gly Ala Val Asp Lys Ser
                695                 700                 705

Tyr Cys Thr Leu Leu Asp Cys Leu Cys Ser Trp Gly Gln Val Gly
                710                 715                 720

His Ile Leu Glu Leu Val Asp Asn Trp Leu Pro Thr Glu His Ala
                725                 730                 735

Gln Ala Lys Ser Asn Thr Ala Ser Lys Gly Arg Val Gln Ile His
                740                 745                 750

Asp Thr Arg Pro Val Lys Pro Glu Leu Ala Leu Val Tyr Ile Glu
                755                 760                 765

Tyr Leu Leu Thr His Pro Lys Asn Arg Glu Cys Leu Leu Ser Ala
                770                 775                 780

Pro Arg Lys Lys Leu Asn His Leu Leu Lys Ala Leu Glu Thr Ser
                785                 790                 795

Lys Ala Asp Leu Glu Ser Leu Leu Gln Thr Pro Gly Gly Lys Pro
                800                 805                 810

Arg Gly Phe Ser Glu Ala Ala Ala Pro Arg Ala Phe Gly Leu His
                815                 820                 825

Cys Arg Leu Ser Ile His Leu Gln His Lys Phe Cys Ser Glu Gly
                830                 835                 840

Lys Val Tyr Leu Ser Met Leu Glu Asp Thr Gly Phe Trp Leu Glu
                845                 850                 855

Ser Lys Ile Leu Ser Phe Ile Gln Asp Gln Glu Glu Asp Tyr Leu
                860                 865                 870

Lys Leu His Arg Val Ile Tyr Gln Gln Ile Ile Gln Thr Tyr Leu
                875                 880                 885

Thr Val Cys Lys Asp Val Val Met Val Gly Leu Gly Asp His Gln
                890                 895                 900

Phe Gln Met Gln Leu Leu Gln Arg Ser Leu Gly Ile Met Gln Thr
                905                 910                 915

Val Lys Gly Phe Phe Tyr Val Ser Leu Leu Leu Asp Ile Leu Lys
                920                 925                 930

Glu Ile Thr Gly Ser Ser Leu Ile Gln Lys Thr Asp Ser Asp Glu
                935                 940                 945

Glu Val Ala Met Leu Leu Asp Thr Val Gln Lys Val Phe Gln Lys
                950                 955                 960
```

```
 Met Leu Glu Cys Ile Ala Arg Ser Phe Arg Lys Gln Pro Glu Glu
                 965                 970                 975

 Gly Leu Arg Leu Leu Tyr Ser Val Gln Arg Pro Leu His Glu Phe
                 980                 985                 990

 Ile Thr Ala Val Gln Ser Arg His Thr Asp Thr Pro Val His Arg
                 995                1000                1005

 Gly Val Leu Ser Thr Leu Ile Ala Gly Pro Val Val Glu Ile Ser
                1010                1015                1020

 His Gln Leu Arg Lys Val Ser Asp Val Glu Glu Leu Thr Pro Pro
                1025                1030                1035

 Glu His Leu Ser Asp Leu Pro Pro Phe Ser Arg Cys Leu Ile Gly
                1040                1045                1050

 Ile Ile Ile Lys Ser Ser Asn Val Val Arg Ser Phe Leu Asp Glu
                1055                1060                1065

 Leu Lys Ala Cys Val Ala Ser Asn Asp Ile Glu Gly Ile Val Cys
                1070                1075                1080

 Leu Thr Ala Ala Val His Ile Ile Leu Val Ile Asn Ala Gly Lys
                1085                1090                1095

 His Lys Ser Ser Lys Val Arg Glu Val Ala Ala Thr Val His Arg
                1100                1105                1110

 Lys Leu Lys Thr Phe Met Glu Ile Thr Leu Glu Glu Asp Ser Ile
                1115                1120                1125

 Glu Arg Phe Leu Tyr Glu Ser Ser Ser Arg Thr Leu Gly Glu Leu
                1130                1135                1140

 Leu Asn Ser


<210> 149
<211> 384
<212> PRT
<213> Homo sapiens

<400> 149
 Met Asn Ser Thr Leu Phe Ser Gln Val Glu Asn His Ser Val His
   1               5                  10                  15

 Ser Asn Phe Ser Glu Lys Asn Ala Gln Leu Leu Ala Phe Glu Asn
                  20                  25                  30

 Asp Asp Cys His Leu Pro Leu Ala Met Ile Phe Thr Leu Ala Leu
                  35                  40                  45

 Ala Tyr Gly Ala Val Ile Ile Leu Gly Val Ser Gly Asn Leu Ala
                  50                  55                  60

 Leu Ile Ile Ile Ile Leu Lys Gln Lys Glu Met Arg Asn Val Thr
                  65                  70                  75

 Asn Ile Leu Ile Val Asn Leu Ser Phe Ser Asp Leu Leu Val Ala
```

```
                 80                  85                  90

Ile Met Cys Leu Pro Phe Thr Phe Val Tyr Thr Leu Met Asp His
                 95                 100                 105

Trp Val Phe Gly Glu Ala Met Cys Lys Leu Asn Pro Phe Val Gln
                110                 115                 120

Cys Val Ser Ile Thr Val Ser Ile Phe Ser Leu Val Leu Ile Ala
                125                 130                 135

Val Glu Arg His Gln Leu Ile Ile Asn Pro Arg Gly Trp Arg Pro
                140                 145                 150

Asn Asn Arg His Ala Tyr Val Gly Ile Ala Val Ile Trp Val Leu
                155                 160                 165

Ala Val Ala Ser Ser Leu Pro Phe Leu Ile Tyr Gln Val Met Thr
                170                 175                 180

Asp Glu Pro Phe Gln Asn Val Thr Leu Asp Ala Tyr Lys Asp Lys
                185                 190                 195

Tyr Val Cys Phe Asp Gln Phe Pro Ser Asp Ser His Arg Leu Ser
                200                 205                 210

Tyr Thr Thr Leu Leu Leu Val Leu Gln Tyr Phe Gly Pro Leu Cys
                215                 220                 225

Phe Ile Phe Ile Cys Tyr Phe Lys Ile Tyr Ile Arg Leu Lys Arg
                230                 235                 240

Arg Asn Asn Met Met Asp Lys Met Arg Asp Asn Lys Tyr Arg Ser
                245                 250                 255

Ser Glu Thr Lys Arg Ile Asn Ile Met Leu Leu Ser Ile Val Val
                260                 265                 270

Ala Phe Ala Val Cys Trp Leu Pro Leu Thr Ile Phe Asn Thr Val
                275                 280                 285

Phe Asp Trp Asn His Gln Ile Ile Ala Thr Cys Asn His Asn Leu
                290                 295                 300

Leu Phe Leu Leu Cys His Leu Thr Ala Met Ile Ser Thr Cys Val
                305                 310                 315

Asn Pro Ile Phe Tyr Gly Phe Leu Asn Lys Asn Phe Gln Arg Asp
                320                 325                 330

Leu Gln Phe Phe Phe Asn Phe Cys Asp Phe Arg Ser Arg Asp Asp
                335                 340                 345

Asp Tyr Glu Thr Ile Ala Met Ser Thr Met His Thr Asp Val Ser
                350                 355                 360

Lys Thr Ser Leu Lys Gln Ala Ser Pro Val Ala Phe Lys Lys Ile
                365                 370                 375

Asn Asn Asn Asp Asp Asn Glu Lys Ile
                380
```

<210> 150
<211> 327
<212> PRT
<213> Homo sapiens

<400> 150
Met Val Arg Cys Asp Arg Gly Leu Gln Met Leu Leu Thr Thr Ala
1 5 10 15

Gly Ala Phe Ala Ala Phe Ser Leu Met Ala Ile Ala Ile Gly Thr
20 25 30

Asp Tyr Trp Leu Tyr Ser Ser Ala His Ile Cys Asn Gly Thr Asn
35 40 45

Leu Thr Met Asp Asp Gly Pro Pro Pro Arg Arg Ala Arg Gly Asp
50 55 60

Leu Thr His Ser Gly Leu Trp Arg Val Cys Cys Ile Glu Gly Ile
65 70 75

Tyr Lys Gly His Cys Phe Arg Ile Asn His Phe Pro Glu Asp Asn
80 85 90

Asp Tyr Asp His Asp Ser Ser Glu Tyr Leu Leu Arg Ile Val Arg
95 100 105

Ala Ser Ser Val Phe Pro Ile Leu Ser Thr Ile Leu Leu Leu Leu
110 115 120

Gly Gly Leu Cys Ile Gly Ala Gly Arg Ile Tyr Ser Arg Lys Asn
125 130 135

Asn Ile Val Leu Ser Ala Gly Ile Leu Phe Val Ala Ala Gly Leu
140 145 150

Ser Asn Ile Ile Gly Ile Ile Val Tyr Ile Ser Ser Asn Thr Gly
155 160 165

Asp Pro Ser Asp Lys Arg Asp Glu Asp Lys Lys Asn His Tyr Asn
170 175 180

Tyr Gly Trp Ser Phe Tyr Phe Gly Ala Leu Ser Phe Ile Val Ala
185 190 195

Glu Thr Val Gly Val Leu Ala Val Asn Ile Tyr Ile Glu Lys Asn
200 205 210

Lys Glu Leu Arg Phe Lys Thr Lys Arg Glu Phe Leu Lys Ala Ser
215 220 225

Ser Ser Ser Pro Tyr Ala Arg Met Pro Ser Tyr Arg Tyr Arg Arg
230 235 240

Arg Arg Ser Arg Ser Ser Ser Arg Ser Thr Glu Ala Ser Pro Ser
245 250 255

Arg Asp Val Ser Pro Met Gly Leu Lys Ile Thr Gly Ala Ile Pro
260 265 270

```
 Met Gly Glu Leu Ser Met Tyr Thr Leu Ser Arg Glu Pro Leu Lys
                 275                 280                 285

 Val Thr Thr Ala Ala Ser Tyr Ser Pro Asp Gln Glu Ala Ser Phe
                 290                 295                 300

 Leu Gln Val His Asp Phe Phe Gln Gln Asp Leu Lys Glu Gly Phe
                 305                 310                 315

 His Val Ser Met Leu Asn Arg Arg Thr Thr Pro Val
                 320                 325

<210> 151
<211> 323
<212> PRT
<213> Homo sapiens

<400> 151
 Met Ala Arg Gly Pro Gly Leu Ala Pro Pro Pro Leu Arg Leu Pro
   1               5                  10                  15

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

 Gln Asp Asn Cys Thr Cys Pro Thr Asn Lys Met Thr Val Cys Ser
                  35                  40                  45

 Pro Asp Gly Pro Gly Gly Arg Cys Gln Cys Arg Ala Leu Gly Ser
                  50                  55                  60

 Gly Met Ala Val Asp Cys Ser Thr Leu Thr Ser Lys Cys Leu Leu
                  65                  70                  75

 Leu Lys Ala Arg Met Ser Ala Pro Lys Asn Ala Arg Thr Leu Val
                  80                  85                  90

 Arg Pro Ser Glu His Ala Leu Val Asp Asn Asp Gly Leu Tyr Asp
                  95                 100                 105

 Pro Asp Cys Asp Pro Glu Gly Arg Phe Lys Ala Arg Gln Cys Asn
                 110                 115                 120

 Gln Thr Ser Val Cys Trp Cys Val Asn Ser Val Gly Val Arg Arg
                 125                 130                 135

 Thr Asp Lys Gly Asp Leu Ser Leu Arg Cys Asp Glu Leu Val Arg
                 140                 145                 150

 Thr His His Ile Leu Ile Asp Leu Arg His Arg Pro Thr Ala Gly
                 155                 160                 165

 Ala Phe Asn His Ser Asp Leu Asp Ala Glu Leu Arg Arg Leu Phe
                 170                 175                 180

 Arg Glu Arg Tyr Arg Leu His Pro Lys Phe Val Ala Ala Val His
                 185                 190                 195

 Tyr Glu Gln Pro Thr Ile Gln Ile Glu Leu Arg Gln Asn Thr Ser
                 200                 205                 210

 Gln Lys Ala Ala Gly Glu Val Asp Ile Gly Asp Ala Ala Tyr Tyr
```

```
                215                 220                 225

 Phe Glu Arg Asp Ile Lys Gly Glu Ser Leu Phe Gln Gly Arg Gly
                230                 235                 240

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

 Thr Leu Ile Tyr Tyr Leu Asp Glu Ile Pro Pro Lys Phe Ser Met
                260                 265                 270

 Lys Arg Leu Thr Ala Gly Leu Ile Ala Val Ile Val Val Val Val
                275                 280                 285

 Val Ala Leu Val Ala Gly Met Ala Val Leu Val Ile Thr Asn Arg
                290                 295                 300

 Arg Lys Ser Gly Lys Tyr Lys Lys Val Glu Ile Lys Glu Leu Gly
                305                 310                 315

 Glu Leu Arg Lys Glu Pro Ser Leu
                320

<210> 152
<211> 207
<212> PRT
<213> Homo sapiens

<400> 152
 Met Gly Lys Cys Ser Gly Arg Cys Thr Leu Val Ala Phe Cys Cys
   1               5                  10                  15

 Leu Gln Leu Val Ala Ala Leu Glu Arg Gln Ile Phe Asp Phe Leu
                20                  25                  30

 Gly Tyr Gln Trp Ala Pro Ile Leu Ala Asn Phe Leu His Ile Met
                35                  40                  45

 Ala Val Ile Leu Gly Ile Phe Gly Thr Val Gln Tyr Arg Ser Arg
                50                  55                  60

 Tyr Leu Ile Leu Tyr Ala Ala Trp Leu Val Leu Trp Val Gly Trp
                65                  70                  75

 Asn Ala Phe Ile Ile Cys Phe Tyr Leu Glu Val Gly Gln Leu Ser
                80                  85                  90

 Gln Asp Arg Asp Phe Ile Met Thr Phe Asn Thr Ser Leu His Arg
                95                 100                 105

 Ser Trp Trp Met Glu Asn Gly Pro Gly Cys Leu Val Thr Pro Val
                110                 115                 120

 Leu Asn Ser Arg Leu Ala Leu Glu Asp His His Val Ile Ser Val
                125                 130                 135

 Thr Gly Cys Leu Leu Asp Tyr Pro Tyr Ile Glu Ala Leu Ser Ser
                140                 145                 150

 Ala Leu Gln Ile Phe Leu Ala Leu Phe Gly Phe Val Phe Ala Cys
                155                 160                 165
```

```
 Tyr Val Ser Lys Val Phe Leu Glu Glu Glu Asp Ser Phe Asp Phe
                 170                 175                 180

 Ile Gly Gly Phe Asp Ser Tyr Gly Tyr Gln Ala Pro Gln Lys Thr
                 185                 190                 195

 Ser His Leu Gln Leu Gln Pro Leu Tyr Thr Ser Gly
                 200                 205

<210> 153
<211> 581
<212> PRT
<213> Homo sapiens

<400> 153
 Met Pro Ala Pro Arg Ala Arg Glu Gln Pro Arg Val Pro Gly Glu
   1               5                  10                  15

 Arg Gln Pro Leu Leu Pro Arg Gly Ala Arg Gly Pro Arg Arg Trp
                  20                  25                  30

 Arg Arg Ala Ala Gly Ala Ala Val Leu Leu Val Glu Met Leu Glu
                  35                  40                  45

 Arg Ala Ala Phe Phe Gly Val Thr Ala Asn Leu Val Leu Tyr Leu
                  50                  55                  60

 Asn Ser Thr Asn Phe Asn Trp Thr Gly Glu Gln Ala Thr Arg Ala
                  65                  70                  75

 Ala Leu Val Phe Leu Gly Ala Ser Tyr Leu Leu Ala Pro Val Gly
                  80                  85                  90

 Gly Trp Leu Ala Asp Val Tyr Leu Gly Arg Tyr Arg Ala Val Ala
                  95                 100                 105

 Leu Ser Leu Leu Leu Tyr Leu Ala Ala Ser Gly Leu Leu Pro Ala
                 110                 115                 120

 Thr Ala Phe Pro Asp Gly Arg Ser Ser Phe Cys Gly Glu Met Pro
                 125                 130                 135

 Ala Ser Pro Leu Gly Pro Ala Cys Pro Ser Ala Gly Cys Pro Arg
                 140                 145                 150

 Ser Ser Pro Ser Pro Tyr Cys Ala Pro Val Leu Tyr Ala Gly Leu
                 155                 160                 165

 Leu Leu Leu Gly Leu Ala Ala Ser Ser Val Arg Ser Asn Leu Thr
                 170                 175                 180

 Ser Phe Gly Ala Asp Gln Val Met Asp Leu Gly Arg Asp Ala Thr
                 185                 190                 195

 Arg Arg Phe Phe Asn Trp Phe Tyr Trp Ser Ile Asn Leu Gly Ala
                 200                 205                 210

 Val Leu Ser Leu Leu Val Val Ala Phe Ile Gln Gln Asn Ile Ser
                 215                 220                 225
```

```
Phe Leu Leu Gly Tyr Ser Ile Pro Val Gly Cys Val Gly Leu Ala
                230                 235                 240

Phe Phe Ile Phe Leu Phe Ala Thr Pro Val Phe Ile Thr Lys Pro
                245                 250                 255

Pro Met Gly Ser Gln Val Ser Ser Met Leu Lys Leu Ala Leu Gln
                260                 265                 270

Asn Cys Cys Pro Gln Leu Trp Gln Arg His Ser Ala Arg Asp Arg
                275                 280                 285

Gln Cys Ala Arg Val Leu Ala Asp Glu Arg Ser Pro Gln Pro Gly
                290                 295                 300

Ala Ser Pro Gln Glu Asp Ile Ala Asn Phe Gln Val Leu Val Lys
                305                 310                 315

Ile Leu Pro Val Met Val Thr Leu Val Pro Tyr Trp Met Val Tyr
                320                 325                 330

Phe Gln Met Gln Ser Thr Tyr Val Leu Gln Gly Leu His Leu His
                335                 340                 345

Ile Pro Asn Ile Phe Pro Ala Asn Pro Ala Asn Ile Ser Val Ala
                350                 355                 360

Leu Arg Ala Gln Gly Ser Ser Tyr Thr Ile Pro Glu Ala Trp Leu
                365                 370                 375

Leu Leu Ala Asn Val Val Val Val Leu Ile Leu Val Pro Leu Lys
                380                 385                 390

Asp Arg Leu Ile Asp Pro Leu Leu Leu Arg Cys Lys Leu Leu Pro
                395                 400                 405

Ser Ala Leu Gln Lys Met Ala Leu Gly Met Phe Phe Gly Phe Thr
                410                 415                 420

Ser Val Ile Val Ala Gly Val Leu Glu Met Glu Arg Leu His Tyr
                425                 430                 435

Ile His His Asn Glu Thr Val Ser Gln Gln Ile Gly Glu Val Leu
                440                 445                 450

Tyr Asn Ala Ala Pro Leu Ser Ile Trp Trp Gln Ile Pro Gln Tyr
                455                 460                 465

Leu Leu Ile Gly Ile Ser Glu Ile Phe Ala Ser Ile Pro Gly Leu
                470                 475                 480

Glu Phe Ala Tyr Ser Glu Ala Pro Arg Ser Met Gln Gly Ala Ile
                485                 490                 495

Met Gly Ile Phe Phe Cys Leu Ser Gly Val Gly Ser Leu Leu Gly
                500                 505                 510

Ser Ser Leu Val Ala Leu Leu Ser Leu Pro Gly Gly Trp Leu His
                515                 520                 525

Cys Pro Lys Asp Phe Gly Asn Ile Asn Asn Cys Arg Met Asp Leu
```

```
                530                 535                 540

 Tyr Phe Phe Leu Leu Ala Gly Ile Gln Ala Val Thr Ala Leu Leu
                545                 550                 555

 Phe Val Trp Ile Ala Gly Arg Tyr Glu Arg Ala Ser Gln Gly Pro
                560                 565                 570

 Ala Ser His Ser Arg Phe Ser Arg Asp Arg Gly
                575                 580

<210> 154
<211> 1266
<212> PRT
<213> Homo sapiens

<400> 154
 Met Trp Val Asn Pro Glu Glu Val Leu Leu Ala Asn Ala Leu Trp
   1               5                  10                  15

 Ile Thr Glu Arg Ala Asn Pro Tyr Phe Ile Leu Gln Arg Arg Lys
                20                  25                  30

 Gly His Ala Gly Asp Gly Gly Gly Gly Gly Gly Leu Ala Gly Leu
                35                  40                  45

 Leu Val Gly Thr Leu Asp Val Val Leu Asp Ser Ser Ala Arg Val
                50                  55                  60

 Ala Pro Tyr Arg Ile Leu Tyr Gln Thr Pro Asp Ser Leu Val Tyr
                65                  70                  75

 Trp Thr Ile Ala Cys Gly Gly Ser Arg Lys Glu Ile Thr Glu His
                80                  85                  90

 Trp Glu Trp Leu Glu Gln Asn Leu Leu Gln Thr Leu Ser Ile Phe
                95                 100                 105

 Glu Asn Glu Asn Asp Ile Thr Thr Phe Val Arg Gly Lys Ile Gln
               110                 115                 120

 Gly Ile Ile Ala Glu Tyr Asn Lys Ile Asn Asp Val Lys Glu Asp
               125                 130                 135

 Asp Asp Thr Glu Lys Phe Lys Glu Ala Ile Val Lys Phe His Arg
               140                 145                 150

 Leu Phe Gly Met Pro Glu Glu Glu Lys Leu Val Asn Tyr Tyr Ser
               155                 160                 165

 Cys Ser Tyr Trp Lys Gly Lys Val Pro Arg Gln Gly Trp Met Tyr
               170                 175                 180

 Leu Ser Ile Asn His Leu Cys Phe Tyr Ser Phe Leu Met Gly Arg
               185                 190                 195

 Glu Ala Lys Leu Val Ile Arg Trp Val Asp Ile Thr Gln Leu Glu
               200                 205                 210

 Lys Asn Ala Thr Leu Leu Leu Pro Asp Val Ile Lys Val Ser Thr
               215                 220                 225
```

```
Arg Ser Ser Glu His Phe Phe Ser Val Phe Leu Asn Ile Asn Glu
                230                 235                 240

Thr Phe Lys Leu Met Glu Gln Leu Ala Asn Ile Ala Met Arg Gln
                245                 250                 255

Leu Leu Asp Asn Glu Gly Phe Glu Gln Asp Arg Ser Leu Pro Lys
                260                 265                 270

Leu Lys Arg Lys Ser Pro Lys Lys Val Ser Ala Leu Lys Arg Asp
                275                 280                 285

Leu Asp Ala Arg Ala Lys Ser Glu Arg Tyr Arg Ala Leu Phe Arg
                290                 295                 300

Leu Pro Lys Asp Glu Lys Leu Asp Gly His Thr Asp Cys Thr Leu
                305                 310                 315

Trp Thr Pro Phe Asn Lys Met His Ile Leu Gly Gln Met Phe Val
                320                 325                 330

Ser Thr Asn Tyr Ile Cys Phe Thr Ser Lys Glu Glu Asn Leu Cys
                335                 340                 345

Ser Leu Ile Ile Pro Leu Arg Glu Val Thr Ile Val Glu Lys Ala
                350                 355                 360

Asp Ser Ser Ser Val Leu Pro Ser Pro Leu Ser Ile Ser Thr Arg
                365                 370                 375

Asn Arg Met Thr Phe Leu Phe Ala Asn Leu Lys Asp Arg Asp Phe
                380                 385                 390

Leu Val Gln Arg Ile Ser Asp Phe Leu Gln Gln Thr Thr Ser Lys
                395                 400                 405

Ile Tyr Ser Asp Lys Glu Phe Ala Gly Ser Tyr Asn Ser Ser Asp
                410                 415                 420

Asp Glu Val Tyr Ser Arg Pro Ser Ser Leu Val Ser Ser Ser Pro
                425                 430                 435

Gln Arg Ser Thr Ser Ser Asp Ala Asp Gly Glu Arg Gln Phe Asn
                440                 445                 450

Leu Asn Gly Asn Ser Val Pro Thr Ala Thr Gln Thr Leu Met Thr
                455                 460                 465

Met Tyr Arg Arg Arg Ser Pro Glu Glu Phe Asn Pro Lys Leu Ala
                470                 475                 480

Lys Glu Phe Leu Lys Glu Gln Ala Trp Lys Ile His Phe Ala Glu
                485                 490                 495

Tyr Gly Gln Gly Ile Cys Met Tyr Arg Thr Glu Lys Thr Arg Glu
                500                 505                 510

Leu Val Leu Lys Gly Ile Pro Glu Ser Met Arg Gly Glu Leu Trp
                515                 520                 525
```

```
Leu Leu Leu Ser Gly Ala Ile Asn Glu Lys Ala Thr His Pro Gly
                530                 535                 540

Tyr Tyr Glu Asp Leu Val Glu Lys Ser Met Gly Lys Tyr Asn Leu
                545                 550                 555

Ala Thr Glu Glu Ile Glu Arg Asp Leu His Arg Ser Leu Pro Glu
                560                 565                 570

His Pro Ala Phe Gln Asn Glu Met Gly Ile Ala Ala Leu Arg Arg
                575                 580                 585

Val Leu Thr Ala Tyr Ala Phe Arg Asn Pro Asn Ile Gly Tyr Cys
                590                 595                 600

Gln Ala Met Asn Ile Val Thr Ser Val Leu Leu Leu Tyr Ala Lys
                605                 610                 615

Glu Glu Glu Ala Phe Trp Leu Leu Val Ala Leu Cys Glu Arg Met
                620                 625                 630

Leu Pro Asp Tyr Tyr Asn Thr Arg Val Val Gly Ala Leu Val Asp
                635                 640                 645

Gln Gly Val Phe Glu Glu Leu Ala Arg Asp Tyr Val Pro Gln Leu
                650                 655                 660

Tyr Asp Cys Met Gln Asp Leu Gly Val Ile Ser Thr Ile Ser Leu
                665                 670                 675

Ser Trp Phe Leu Thr Leu Phe Leu Ser Val Met Pro Phe Glu Ser
                680                 685                 690

Ala Val Val Val Val Asp Cys Phe Phe Tyr Glu Gly Ile Lys Val
                695                 700                 705

Ile Phe Gln Leu Ala Leu Ala Val Leu Asp Ala Asn Val Asp Lys
                710                 715                 720

Leu Leu Asn Cys Lys Asp Asp Gly Glu Ala Met Thr Val Leu Gly
                725                 730                 735

Arg Tyr Leu Asp Ser Val Thr Asn Lys Asp Ser Thr Leu Pro Pro
                740                 745                 750

Ile Pro His Leu His Ser Leu Leu Ser Asp Asp Val Glu Pro Tyr
                755                 760                 765

Pro Glu Val Asp Ile Phe Arg Leu Ile Arg Thr Ser Tyr Glu Lys
                770                 775                 780

Phe Gly Thr Ile Arg Ala Asp Leu Ile Glu Gln Met Arg Phe Lys
                785                 790                 795

Gln Arg Leu Lys Val Ile Gln Thr Leu Glu Asp Thr Thr Lys Arg
                800                 805                 810

Asn Val Val Arg Thr Ile Val Thr Glu Thr Ser Phe Thr Ile Asp
                815                 820                 825

Glu Leu Glu Glu Leu Tyr Ala Leu Phe Lys Ala Glu His Leu Thr
```

```
                830                 835                 840

Ser Cys Tyr Trp Gly Gly Ser Ser Asn Ala Leu Asp Arg His Asp
                845                 850                 855

Pro Ser Leu Pro Tyr Leu Glu Gln Tyr Arg Ile Asp Phe Glu Gln
                860                 865                 870

Phe Lys Gly Met Phe Ala Leu Leu Phe Pro Trp Ala Cys Gly Thr
                875                 880                 885

His Ser Asp Val Leu Ala Ser Arg Leu Phe Gln Leu Leu Asp Glu
                890                 895                 900

Asn Gly Asp Ser Leu Ile Asn Phe Arg Glu Phe Val Ser Gly Leu
                905                 910                 915

Ser Ala Ala Cys His Gly Asp Leu Thr Glu Lys Leu Lys Leu Leu
                920                 925                 930

Tyr Lys Met His Val Leu Pro Glu Pro Ser Ser Asp Gln Asp Glu
                935                 940                 945

Pro Asp Ser Ala Phe Glu Ala Thr Gln Tyr Phe Phe Glu Asp Ile
                950                 955                 960

Thr Pro Glu Cys Thr His Val Val Gly Leu Asp Ser Arg Ser Lys
                965                 970                 975

Gln Gly Ala Asp Asp Gly Phe Val Thr Val Ser Leu Lys Pro Asp
                980                 985                 990

Lys Gly Lys Arg Ala Asn Ser Gln Glu Asn Arg Asn Tyr Leu Arg
                995                1000                1005

Leu Trp Thr Pro Glu Asn Lys Ser Lys Ser Lys Asn Ala Lys Asp
               1010                1015                1020

Leu Pro Lys Leu Asn Gln Gly Gln Phe Ile Glu Leu Cys Lys Thr
               1025                1030                1035

Met Tyr Asn Met Phe Ser Glu Asp Pro Asn Glu Gln Glu Leu Tyr
               1040                1045                1050

His Ala Thr Ala Ala Val Thr Ser Leu Leu Leu Glu Ile Gly Glu
               1055                1060                1065

Val Gly Lys Leu Phe Val Ala Gln Pro Ala Lys Glu Gly Gly Ser
               1070                1075                1080

Gly Gly Ser Gly Pro Ser Cys His Gln Gly Ile Pro Gly Val Leu
               1085                1090                1095

Phe Pro Lys Lys Gly Pro Gly Gln Pro Tyr Val Val Glu Ser Val
               1100                1105                1110

Glu Pro Leu Pro Ala Ser Leu Ala Pro Asp Ser Glu Glu His Ser
               1115                1120                1125

Leu Gly Gly Gln Met Glu Asp Ile Lys Leu Glu Asp Ser Ser Pro
               1130                1135                1140
```

Arg Asp Asn Gly Ala Cys Ser Ser Met Leu Ile Ser Asp Asp Asp
1145 1150 1155

Thr Lys Asp Asp Ser Ser Met Ser Ser Tyr Ser Val Leu Ser Ala
1160 1165 1170

Gly Ser His Glu Glu Asp Lys Leu His Cys Glu Asp Ile Gly Glu
1175 1180 1185

Asp Thr Val Leu Val Arg Ser Gly Gln Gly Thr Ala Ala Leu Pro
1190 1195 1200

Arg Ser Thr Ser Leu Asp Arg Asp Trp Ala Ile Thr Phe Glu Gln
1205 1210 1215

Phe Leu Ala Ser Leu Leu Thr Glu Pro Ala Leu Val Lys Tyr Phe
1220 1225 1230

Asp Lys Pro Val Cys Met Met Ala Arg Ile Thr Ser Ala Lys Asn
1235 1240 1245

Ile Arg Met Met Gly Lys Pro Leu Thr Ser Ala Ser Asp Tyr Glu
1250 1255 1260

Ile Ser Ala Met Ser Gly
1265

<210> 155
<211> 21
<212> DNA
<213> Homo sapiens

<400> 155
aaggaagagg cgcgggagaa g 21

<210> 156
<211> 21
<212> DNA
<213> Homo sapiens

<400> 156
ggaagaggcg cgggagaagt t 21

<210> 157
<211> 21
<212> DNA
<213> Homo sapiens

<400> 157
cuucucccgc gccucuucct t 21

<210> 158
<211> 21
<212> DNA
<213> Homo sapiens

<400> 158
cuggacuucc agaagaacat t 21

<210> 159

<211> 18
<212> DNA
<213> Homo sapiens

<400> 159
uguucuggaa guccagtt 18