(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 July 2003 (03.07.2003)

PCT

(10) International Publication Number
WO 03/053364 A2

(51) International Patent Classification[7]: A61K

(21) International Application Number: PCT/US02/41011

(22) International Filing Date:
19 December 2002 (19.12.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/341,953 19 December 2001 (19.12.2001) US

(71) Applicant *(for all designated States except US)*: **MILLENNIUM PHARMACEUTICALS, INC.** [US/US]; 75 Sidney Street, Cambridge, MA 01239 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SILOS-SANTIAGO, Inmaculada** [ES/US]; 2123 De Mayo Road, Del Mar, CA 92014 (US). **ROSENFELD, Julie, Beth** [US/US]; 9 Red Fox Run, Sharon, MA 02067 (US).

(74) Agent: **BOSSONE, Steven, A.**; Millennium Pharmaceuticals, Inc., 75 Sidney Street, Cambridge, MA 02139 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/053364 A2

(54) Title: METHODS AND COMPOSITIONS IN TREATING PAIN AND PAINFUL DISORDERS USING 1465,1587, 2146, 2207, 32838, 336 AND 52908

(57) Abstract: The present invention relates to methods for the diagnosis and treatment of pain or painful disorders. Specifically, the present invention identifies the differential expression of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 genes in tissues relating to pain sensation, relative to their expression in normal, or non-painful disease states, and/or in response to manipulations relevant to pain. The present invention describes methods for the diagnostic evaluation and prognosis of various cardiovascular diseases, and for the identification of subjects exhibiting a predisposition to such conditions. The invention also provides methods for identifying a compound capable of modulating pain or painful disorders. The present invention also provides methods for the identification and therapeutic use of compounds as treatments of pain and painful disorders.

METHODS AND COMPOSITIONS IN TREATING PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207, 32838, 336 AND 52908

The present application claims priority from U.S. provisional patent application serial no. 60/341,953, filed on December 19, 2001. The above application is expressly incorporated by reference.

The sensation of pain can be categorized into two types, peripheral and central pain. Peripheral pain can be classified into three broad areas, nociceptive pain, inflammatory pain and neuropathic pain. Nociceptive pain is also referred to as physiological pain and serves as a defense mechanism throughout the animal kingdom. Inflammatory pain, arising from severe wounds and/or associated with inflammatory infiltrates, can be well controlled by non-steroidal anti-inflammatory drugs (NSAID)-like drugs, steroids and opiates. However, the etiology and management of neuropathic pain is not well understood. Neuropathic pain is thought to arise from inherent defects in sensory and as a consequence in sympathetic neurons and can be secondary to trauma.

Peripheral pain is mediated by two types of primary sensory neuron classes, the Ad- and C-fibers, whose cell bodies lie within the dorsal root ganglion. Although the mechanisms of generation of neuropathic pain are poorly understood it is clear that several factors influence the perception and transmission of the painful stimulus, namely, alterations in chemical environment, ectopic generation of sensory neuron firing and sympathetic discharge. Some of the most common syndromes associated with neuropathic pain arise from destruction of small sensory fibers (or possibly the alteration in ratios of small to large fibers) as it is common in post-traumatic situations. Other etiologies of pain arise from small fiber damage due to diabetic neuropathy, drug induced damage (chemotherapy drugs), alcoholism, damage due to cancer, and a variety of hereditary small- and large-fiber neuropathies. We rationalize that targets derived from the peripheral nervous system may be of strategic benefit in that candidate compounds do not need to cross the blood-brain barrier, they can act on the initiation site of pain without inducing central side effects.

It has long been established that central mechanisms are involved in the perception and modulation of pain. Electrical stimulation of the periaqueductal gray (PAG) area produces analgesia without loss of other sensory modalities. Descending pain pathways emanating from PAG and the nucleus raphe magnus impinge on dorsal spinal cord regions where primary nociceptive afferents terminate. Also, stimulation of regions such as the

paragigantocellularis nucleus in the medulla oblongata result in analgesia. Finally, opiate receptors, when stimulated by opioid alkaloids and opioid peptides, mediate analgesia and these sites are located in key "pain centers" within the brain including PAG, thalamic nuclei and cortical regions. Identification of genes in these CNS regions and the spinal thalamic tract from animal models of pain may elucidate important targets for pain modulation.

The present invention provides methods and compositions for the diagnosis and treatment of a subject experiencing pain or suffering from a painful disorders. Preferably, the subject is a human, e.g., a patient with pain or a pain-associated disorder disclosed herein. For example, the subject can be a patient with pain elicited from tissue injury, e.g., inflammation, infection, ischemia; pain associated with musculoskeletal disorders, e.g., joint pain; tooth pain; headaches, e.g., migrane; pain associated with surgery; pain related to inflammation, e.g., irritable bowel syndrome; or chest pain. The subject can be a patient with complex regional pain syndrome (CRPS), reflex sympathetic dystrophy (RSD), causalgia, neuralgia, central pain and dysesthesia syndrome, carotidynia, neurogenic pain, refractory cervicobrachial pain syndrome, myofascial pain syndrome, craniomandibular pain dysfunction syndrome, chronic idiopathic pain syndrome, Costen's pain-dysfunction, acute chest pain syndrome, gynecologic pain syndrome, patellofemoral pain syndrome, anterior knee pain syndrome, recurrent abdominal pain in children, colic, low back pain syndrome, neuropathic pain, phantom pain from amputation, phantom tooth pain, or pain asymbolia. The subject can be a cancer patient, e.g., a patient with brain cancer, bone cancer, or prostate cancer. In other embodiments, the subject is a non-human animal, e.g., an experimental animal, e.g., an arthritic rat model of chronic pain, a chronic constriction injury (CCI) rat model of neuropathic pain, or a rat model of unilateral inflammatory pain by intraplantar injection of Freund's complete adjuvant (FCA).

"Treatment", as used herein, is defined as the application or administration of a therapeutic agent to a patient, or application or administration of a therapeutic agent to an isolated tissue or cell line from a patient, who has a disease or disorder, a symptom of disease or disorder or a predisposition toward a disease or disorder, with the purpose of curing, healing, alleviating, relieving, altering, remedying, ameliorating, improving or affecting the disease or disorder, the symptoms of disease or disorder or the predisposition toward a disease or disorder. A therapeutic agent includes, but is not limited to, the small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides described herein.

The present invention is based, at least in part, on the discovery that nucleic acid and protein molecules, (described infra), are differentially expressed in animal models of pain and in peripheral and central nervous system tissues known to be associated with pain (e.g. dorsal root ganglion (DRG)). The modulators of the molecules of the present invention, identified according to the methods of the invention can be used to modulate (*e.g.*, inhibit, treat, or prevent) pain and painful conditions.

"Differential expression", as used herein, includes both quantitative as well as qualitative differences in the temporal and/or tissue expression pattern of a gene. Thus, a differentially expressed gene may have its expression activated or inactivated in normal versus painful disease conditions (for example, in an experimental pain model system such as in an animal model for pain). The degree to which expression differs in normal versus treated or control versus experimental states need only be large enough to be visualized via standard characterization techniques, *e.g.*, quantitative PCR, Northern analysis, subtractive hybridization. The expression pattern of a differentially expressed gene may be used as part of a prognostic or diagnostic, evaluation, or may be used in methods for identifying compounds useful for the treatment of pain and painful disorders. In addition, a differentially expressed gene involved in pain may represent a target gene such that modulation of the level of target gene expression or of target gene product activity may act to ameliorate a painful disease condition. Compounds that modulate target gene expression or activity of the target gene product can be used in the treatment of pain or painful conditions. Although the genes described herein may be differentially expressed with respect to pain, and/or their products may interact with gene products important to pain, the genes may also be involved in mechanisms important to additional cell processes.

Molecules of the Present Invention

Molecules of the present invention include, but are not limited to ion channels (eg. Potassium channels), transporters (e.g. amino acid transporters), receptors (e.g. G protein coupled receptors) and enzymes (e.g. kinases)

Transmembrane ion channel proteins that selectively mediate the conductance of sodium, potassium, calcium and chloride ions directly modulate the electrical activity of sensory neurons and are, thus, important in nociception. In particular, potassium channels are main players in regulating the frequency and pattern of neuronal firing. The expression and peak currents of potassium channels has been shown to be regulated after different models of inflammatory and chronic pain. Additionally, calcium ions serve important intracellular

signaling roles including modulation of other ion channels and regulation of protein kinases and other enzymatic activity. As cell surface proteins with established three-dimensional structures and modes of action, the pore-forming alpha subunits of ion channels make ideal drug targets. In addition to alpha subunits, these channels may consist of beta subunits and other interacting proteins which modulate channel activity and are good targets for pharmacological manipulation of the channels. Therefore, ion channels are useful in treating pain and painful conditions.

Endogenous soluble factors mediate pain sensation by binding to specific transmembrane receptors either on the peripheral terminals of nociceptive neurons or on central neurons receiving input from these nociceptors. These soluble factors include, but are not limited to serotonine, histamine, bradykinin, tachykinins (substance P and neurokinin A), opioids, eicosanoids (leukotrienes, prostaglandins, thromboxanes), purines, excitatory amino acids and different proteins. In addition a growing body of evidence, including clinical trials in man, indicates that IL-1, TNFα, and members of the neurotrophin family are involved at several stages in the transmission of painful stimuli. Hydrogen ions (protons) may mediate pain associated with inflammation (and also acid taste) by activating vanilloid receptor calcium channels or amiloride-sensitive sodium channels. Additionally, numerous exogenous agents modulate pain by mimicking endogenous soluble factors. For instance the opiate drugs of abuse exert analgesic effects by binding to receptors for the endogenous opioids and capsaicin stimulates pain sensation by binding to vanilloid receptors. The receptors for these soluble factors are linked to several signal transduction mechanisms including tyrosine kinase activity (e.g. neurotrophin receptors), recruitment of cytoplasmic tyrosine kinases (e.g. cytokine receptors for TNFα and IL-1), ion channel opening, and G-protein coupled receptors. These cell surface receptors are ideal drug targets due to their transmembrane location, and the goal is to discover G-protein coupling receptors with known ligands or with surrogate ligands that may be important players in regulating pain mechanisms.

Intracellular kinases such as protein kinase A and protein kinase C are involved in the response to pain in sensory neurons. Similarly, enzymes such as cyclooxygenase(s) and thromboxane synthetase are know to be critical in the production of prostaglandins, leukotrienes and thromboxanes. Although these particular targets may be more important in inflammatory pain, the role of this gene family in long term or neuropathic pain is of importance.

Gene ID 1465

The human 1465 sequence (SEQ ID NO:1), (GI:971256, known also as G protein-coupled receptor kinase GRK4) which is approximately 2113 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 1737 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ ID NO:1, SEQ ID NO:3). The coding sequence encodes a 578 amino acid protein (SEQ ID NO:2) (GI:971257).

As assessed by TaqMan® analysis, 1465 was expressed in the brain and dorsal root ganglion (DRG). Therefore, it is involved in pain disorders and would be a potential target to discover therapeutics directed toward the treatment of pain and painful disorders.

Gene ID 1587

The human 1587 sequence (SEQ ID NO:4), (GI:297101), known also as serine/threonine protein kinase PCTAORE-3) which is approximately 1242 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 1143 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ ID NO:4, SEQ ID NO:6). The coding sequence encodes a 380 amino acid protein (SEQ ID NO:5) (GI:297102).

As assessed by TaqMan® analysis, 1587 shows significant expression in brain and spinal cord samples as well as heart, DRG and kidney tissues in both human and rat TaqMan® panels. Morphine experiments demonstrated a marked decrease in 1587 at 6 hours and 24 hours in the dorsal horn of the spinal cord (DH), and at 2 and 11 days in the DRG. In animal pain model systems, 1587 displayed significant down regulation in morphine and capsaicin treated animals. ISH experiments have confirmed the expression of 1587 in brain and DRG tissue samples as well as in subpopulations of glial cells. Therefore, 1587 is involved in pain response and would be a potential target to discover therapeutics directed toward the treatment of pain and painful disorders.

Gene ID 2146

The human 2146 sequence (SEQ ID NO:7), (GI:988304) (known also as protein tyrosine kinase PYK2) which is approximately 4151 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 3030 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ

ID NO:7, SEQ ID NO:9). The coding sequence encodes a 1009 amino acid protein (SEQ ID NO:8) (GI:988305).

As assessed by TaqMan® analysis, 2146 was expressed at high levels in brain, DRG, spleen and peripheral blood cells. ISH experiments conducted on monkey and rat samples have confirmed the expression of 2146 in the spinal cord and DRG. In the spinal cord 2146 is expressed only in the most superficial laminae, the region involved in nociception. In the monkey and rat DRG samples, 2146 expression is observed in a very restricted subpopulation of neurons, mainly of small and intermediate diameter. Therefore, 2146 is involved in pain disorders and would be a potential target to discover therapeutics directed toward the treatment of pain and painful disorders.

Gene ID 2207

The human 2207 sequence (SEQ ID NO:10), (known also as a protein kinase) which is approximately 6574 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 6162 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ ID NO:10, SEQ ID NO:12). The coding sequence encodes a 2053 amino acid protein (SEQ ID NO:11).

As assessed by TaqMan® analysis, 2207 mRNA was highly expressed in the brain, peripheral blood cells and DRG, and at lower levels in the spinal cord. 2207 is also upregulated in the capsaicin model of pain. ISH experiments conducted on monkey and rat samples confirmed the expression of 2207 in the spinal cord and DRG. In the brain, the highest levels of expression are in the sensory thalamus. In monkey and rat DRG samples, expression of 2207 is restricted to a subpopulation of neurons of all sizes, including neurons involved in pain processing. Therefore, 2207 is involved in pain disorders and would be a potential target to discover therapeutics directed toward the treatment of pain and painful disorders.

Gene ID 32838

The human 32838 sequence (SEQ ID NO:13), (GI:4539524) (known also as NAALADase II) which is approximately 3152 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 2223 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ ID NO:13, SEQ ID NO:15). The coding sequence encodes a 740 amino acid protein (SEQ ID NO:14) (GI:4535925).

As assessed by TaqMan® analysis, 32838 mRNA was restircted to expression in the brain cortex, pituitary and hypothalamus. It was also upregulated at 6 and 24 hours in the DH of animals treated with morphine. Therefore, 32838 is involved in pain disorders and would be a potential target to discover therapeutics directed toward the treatment of pain and painful disorders.

Gene ID 336

The human 336 sequence (SEQ ID NO:16), (GI:2246432) (known also as CCR6 chemokine receptor CMKBR6) which is approximately 3693 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 1221 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ ID NO:16, SEQ ID NO:18). The coding sequence encodes a 406 amino acid protein (SEQ ID NO:17) (GI:2251211).

As assessed by TaqMan® analysis, 336 mRNA was expressed in brain cortex, indicating a role in pain perception. Therefore, 336 is a potential target to discover therapeutics useful in treating pain and painful disorders.

Gene ID 52908

The human 52908 sequence (SEQ ID NO:19), (known also as a potassium channel) which is approximately 3164 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 2877 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ ID NO:19, SEQ ID NO:21). The coding sequence encodes a 958 amino acid protein (SEQ ID NO:20).

As assessed by TaqMan® analysis, 52908 mRNA expression was restricted to brain cortex, hypothalamus, spinal cord and DRG. ISH experiments confirmed the expression of 2207 in DRG and spinal cord. In the spinal cord, 52908 is expressed in a sub-population of neurons in laminae I, II and V. Also, some expression is observed around the central canal, lamina X. High levels of 52908 expression was also observed in a sub-population of nociceptive DRG neurons. This neuronal population corresponds to that of small and intermediate diameter. Therefore, 52908 is involved in pain and would be a potential target to discover therapeutics directed toward the treatment of pain and painful disorders.

Various aspects of the invention are described in further detail in the following subsections:

I. Screening Assays:

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules (organic or inorganic) or other drugs) which bind to 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins, have a stimulatory or inhibitory effect on, for example, 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression or 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 substrate. Compounds identified using the assays described herein may be useful for treating pain and painful conditions.

These assays are designed to identify compounds that bind to a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, bind to other intracellular or extracellular proteins that interact with a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, and interfere with the interaction of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein with other intercellular or extracellular proteins. For example, in the case of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, which is a transmembrane receptor-type protein, such techniques can identify ligands for such a receptor. A 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein ligand or substrate can, for example, be used to ameliorate pain and painful conditions. Such compounds may include, but are not limited to peptides, antibodies, or small organic or inorganic compounds. Such compounds may also include other cellular proteins.

Compounds identified via assays such as those described herein may be useful, for example, for treating pain and painful conditions. In instances whereby a painful condition results from an overall lower level of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene expression and/or 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein in a cell or tissue, compounds that interact with the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein may include compounds which accentuate or amplify the activity of the bound 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. Such compounds would bring about an effective increase in the level of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein activity, thus ameliorating symptoms.

In other instances, mutations within the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene may cause aberrant types or excessive amounts of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins to be made which have a deleterious effect that leads to a pain. Similarly, physiological conditions may cause an excessive increase in 1465, 1587,

2146, 2207, 32838, 336 OR 52908 gene expression leading pain. In such cases, compounds that bind to a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein may be identified that inhibit the activity of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. Assays for testing the effectiveness of compounds identified by techniques such as those described in this section are discussed herein.

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or polypeptide or biologically active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994). *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate

1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity is determined. Determining the ability of the test compound to modulate 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity can be accomplished by monitoring, for example, intracellular calcium, IP_3, cAMP, or diacylglycerol concentration, the phosphorylation profile of intracellular proteins, cell proliferation and/or migration, gene expression of, for example, cell surface adhesion molecules or genes associated with analgesia, or the activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -regulated transcription factor. The cell can be of mammalian origin, *e.g.*, a neural cell. In one embodiment, compounds that interact with a receptor domain can be screened for their ability to function as ligands, *i.e.*, to bind to the receptor and modulate a signal transduction pathway. Identification of ligands, and measuring the activity of the ligand-receptor complex, leads to the identification of modulators (*e.g.*, antagonists) of this interaction. Such modulators may be useful in the treatment of pain and painful conditions.

The ability of the test compound to modulate 1465, 1587, 2146, 2207, 32838, 336 OR 52908 binding to a substrate or to bind to 1465, 1587, 2146, 2207, 32838, 336 OR 52908 can also be determined. Determining the ability of the test compound to modulate 1465, 1587, 2146, 2207, 32838, 336 OR 52908 binding to a substrate can be accomplished, for example, by coupling the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 substrate with a radioisotope or enzymatic label such that binding of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 substrate to 1465, 1587, 2146, 2207, 32838, 336 OR 52908 can be determined by detecting the labeled 1465, 1587, 2146, 2207, 32838, 336 OR 52908 substrate in a complex. 1465, 1587, 2146, 2207, 32838, 336 OR 52908 could also be coupled with a radioisotope or enzymatic label to monitor the ability of a test compound to modulate 1465, 1587, 2146, 2207, 32838, 336 OR 52908 binding to a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 substrate in a complex. Determining the ability of the test compound to bind 1465, 1587, 2146, 2207, 32838, 336 OR 52908 can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that binding of the compound to 1465, 1587, 2146, 2207, 32838, 336 OR 52908 can be determined by detecting the labeled 1465, 1587, 2146, 2207, 32838, 336 OR 52908 compound in a complex. For example, compounds (*e.g.*, 1465, 1587, 2146, 2207, 32838, 336 OR 52908 ligands or substrates) can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemmission or by scintillation counting. Compounds can further be enzymatically labeled with, for

example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

It is also within the scope of this invention to determine the ability of a compound (*e.g.*, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 ligand or substrate) to interact with 1465, 1587, 2146, 2207, 32838, 336 OR 52908 without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of a compound with 1465, 1587, 2146, 2207, 32838, 336 OR 52908 without the labeling of either the compound or the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 (McConnell, H. M. *et al.* (1992) *Science* 257:1906-1912. As used herein, a "microphysiometer" (*e.g.*, Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a compound and 1465, 1587, 2146, 2207, 32838, 336 OR 52908 .

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 target molecule (*e.g.*, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 substrate) with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 target molecule. Determining the ability of the test compound to modulate the activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 target molecule can be accomplished, for example, by determining the ability of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein to bind to or interact with the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 target molecule.

Determining the ability of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or a biologically active fragment thereof, to bind to or interact with a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 target molecule can be accomplished by one of the methods described above for determining direct binding. In a preferred embodiment, determining the ability of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein to bind to or interact with a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.*, intracellular Ca^{2+}, diacylglycerol, IP_3, cAMP), detecting catalytic/enzymatic activity of the target on an appropriate substrate, detecting the induction of a reporter gene (comprising a target-responsive regulatory element

operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a target-regulated cellular response (*e.g.*, gene expression).

In yet another embodiment, an assay of the present invention is a cell-free assay in which a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or biologically active portion thereof, is contacted with a test compound and the ability of the test compound to bind to the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or biologically active portion thereof is determined. Preferred biologically active portions of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins to be used in assays of the present invention include fragments which participate in interactions with non-1465, 1587, 2146, 2207, 32838, 336 OR 52908 molecules, *e.g.*, fragments with high surface probability scores. Binding of the test compound to the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or biologically active portion thereof with a known compound which binds 1465, 1587, 2146, 2207, 32838, 336 OR 52908 to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, wherein determining the ability of the test compound to interact with a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein comprises determining the ability of the test compound to preferentially bind to 1465, 1587, 2146, 2207, 32838, 336 OR 52908 or biologically active portion thereof as compared to the known compound. Compounds that modulate the interaction of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 with a known target protein may be useful in regulating the activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, especially a mutant 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein.

In another embodiment, the assay is a cell-free assay in which a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or biologically active portion thereof is determined. Determining the ability of the test compound to modulate the activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein can be accomplished, for example, by determining the ability of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein to bind to a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 target molecule by one of the methods described above for determining direct binding. Determining the ability of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908

protein to bind to a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 target molecule can also be accomplished using a technology such as real-time Biomolecular Interaction Analysis (BIA) (Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705). As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In another embodiment, determining the ability of the test compound to modulate the activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein can be accomplished by determining the ability of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein to further modulate the activity of a downstream effector of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 target molecule. For example, the activity of the effector molecule on an appropriate target can be determined or the binding of the effector to an appropriate target can be determined as previously described.

In yet another embodiment, the cell-free assay involves contacting a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or biologically active portion thereof with a known compound which binds the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, wherein determining the ability of the test compound to interact with the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein comprises determining the ability of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein to preferentially bind to or modulate the activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 target molecule.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either 1465, 1587, 2146, 2207, 32838, 336 OR 52908 or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, or interaction of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which

adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/1465, 1587, 2146, 2207, 32838, 336 OR 52908 fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, and the mixture incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or target molecules but which do not interfere with binding of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein to its target molecule can be derivatized to the wells of the plate, and unbound target or 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or target molecule.

In another embodiment, modulators of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA or protein in the cell is determined. The level of expression of 1465, 1587, 2146, 2207,

32838, 336 OR 52908 mRNA or protein in the presence of the candidate compound is compared to the level of expression of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression based on this comparison. For example, when expression of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA or protein expression. Alternatively, when expression of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA or protein expression. The level of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA or protein expression in the cells can be determined by methods described herein for detecting 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA or protein.

In yet another aspect of the invention, the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with 1465, 1587, 2146, 2207, 32838, 336 OR 52908 ("1465, 1587, 2146, 2207, 32838, 336 OR 52908 -binding proteins" or "1465, 1587, 2146, 2207, 32838, 336 OR 52908 -bp") and are involved in 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity. Such 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -binding proteins are also likely to be involved in the propagation of signals by the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins or 1465, 1587, 2146, 2207, 32838, 336 OR 52908 targets as, for example, downstream elements of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -mediated signaling pathway. Alternatively, such 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -binding proteins are likely to be 1465, 1587, 2146, 2207, 32838, 336 OR 52908 inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein is fused to a gene encoding the

DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo,* forming a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein can be confirmed *in vivo*, *e.g.*, in an animal such as an animal model for pain, as described herein.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulating agent, an antisense 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid molecule, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -specific antibody, or a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

Any of the compounds, including but not limited to compounds such as those identified in the foregoing assay systems, may be tested for the ability to ameliorate pain. Cell-based and animal model-based assays for the identification of compounds exhibiting such an ability to ameliorate pain are described herein.

In addition, animal-based models of pain, such as those described herein, may be used to identify compounds capable of treating pain and painful conditions. Such animal models may be used as test substrates for the identification of drugs, pharmaceuticals, therapies, and interventions which may be effective in treating pain. For example, animal models may be exposed to a compound, suspected of exhibiting an ability to treat pain, at a sufficient concentration and for a time sufficient to elicit such an amelioration of pain in the exposed animals. The response of the animals to the exposure may be monitored by assessing the reversal of the symptoms of pain before and after treatment.

With regard to intervention, any treatments which reverse any aspect of pain (i.e. have an analgesic effect) should be considered as candidates for human pain therapeutic intervention. Dosages of test agents may be determined by deriving dose-response curves.

Additionally, gene expression patterns may be utilized to assess the ability of a compound to ameliorate pain. For example, the expression pattern of one or more genes may form part of a "gene expression profile" or “transcriptional profile” which may be then be used in such an assessment. "Gene expression profile" or “transcriptional profile”, as used herein, includes the pattern of mRNA expression obtained for a given tissue or cell type under a given set of conditions. Gene expression profiles may be generated, for example, by utilizing a differential display procedure, Northern analysis and/or RT-PCR. In one embodiment, 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene sequences may be used as probes and/or PCR primers for the generation and corroboration of such gene expression profiles.

Gene expression profiles may be characterized for known states, either cardiovascular disease or normal, within the cell- and/or animal-based model systems. Subsequently, these known gene expression profiles may be compared to ascertain the effect a test compound has to modify such gene expression profiles, and to cause the profile to more closely resemble that of a more desirable profile.

For example, administration of a compound may cause the gene expression profile of a pain disease model system to more closely resemble the control system. Administration of a compound may, alternatively, cause the gene expression profile of a control system to begin to mimic pain or a painful disease state. Such a compound may, for example, be used in further characterizing the compound of interest, or may be used in the generation of additional animal models.

II. Cell- and Animal-Based Model Systems

Described herein are cell- and animal-based systems which act as models for pain. These systems may be used in a variety of applications. For example, the cell- and animal-based model systems may be used to further characterize differentially expressed genes associated with cardiovascular disease, *e.g.*, 1465, 1587, 2146, 2207, 32838, 336 OR 52908 . In addition, animal- and cell-based assays may be used as part of screening strategies designed to identify compounds which are capable of ameliorating pain, as described, below. Thus, the animal- and cell-based models may be used to identify drugs, pharmaceuticals, therapies and interventions which may be effective in treating cardiovascular disease. Furthermore, such animal models may be used to determine the LD50 and the ED50 in animal subjects, and such data can be used to determine the *in vivo* efficacy of potential pain treatments.

A. Animal-Based Systems

Animal-based model systems of pain may include, but are not limited to, non-recombinant and engineered transgenic animals.

Non-recombinant animal models for pain may include, for example, genetic models.

Additionally, animal models exhibiting pain may be engineered by using, for example, 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene sequences described above, in conjunction with techniques for producing transgenic animals that are well known to those of skill in the art. For example, 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene sequences may be introduced into, and overexpressed in, the genome of the animal of interest, or, if endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene sequences are present, they may either be overexpressed or, alternatively, be disrupted in order to underexpress or inactivate 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene expression.

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 sequences have been introduced into their genome or homologous recombinant animals in which endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 sequences

have been altered. Such animals are useful for studying the function and/or activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 and for identifying and/or evaluating modulators of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, and the like. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal used in the methods of the invention can be created by introducing a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The 1465, 1587, 2146, 2207, 32838, 336 OR 52908 cDNA sequence can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of a human 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene, such as a mouse or rat 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene, can be used as a transgene. Alternatively, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene homologue, such as another 1465, 1587, 2146, 2207, 32838, 336 OR 52908 family member, can be isolated based on hybridization to the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 cDNA sequences and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 transgene to direct expression of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan,

B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 transgene in its genome and/or expression of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene. The 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene can be a human gene but more preferably, is a non-human homologue of a human 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene. For example, a rat 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene can be used to construct a homologous recombination nucleic acid molecule, *e.g.*, a vector, suitable for altering an endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene in the mouse genome. In a preferred embodiment, the homologous recombination nucleic acid molecule is designed such that, upon homologous recombination, the endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the homologous recombination nucleic acid molecule can be designed such that, upon homologous recombination, the endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein). In the homologous recombination nucleic acid molecule, the altered portion of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene is flanked at its 5' and 3' ends by additional nucleic acid sequence of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene to allow for homologous recombination to occur between the exogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene carried by the homologous recombination nucleic acid molecule and an endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene in a cell, *e.g.*, an embryonic stem cell. The additional flanking 1465, 1587, 2146, 2207, 32838, 336

OR 52908 nucleic acid sequence is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the homologous recombination nucleic acid molecule (see, *e.g.*, Thomas, K.R. and Capecchi, M. R. (1987) *Cell* 51:503 for a description of homologous recombination vectors). The homologous recombination nucleic acid molecule is introduced into a cell, *e.g.*, an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene has homologously recombined with the endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene are selected (see *e.g.*, Li, E. *et al.* (1992) *Cell* 69:915). The selected cells can then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras (see *e.g.*, Bradley, A. 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. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination nucleic acid molecules, *e.g.*, vectors, or homologous recombinant animals are described further in Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec *et al.*; WO 91/01140 by Smithies *et al.*; WO 92/0968 by Zijlstra *et al.*; and WO 93/04169 by Berns *et al.*

In another embodiment, transgenic non-human animals for use in the methods of the invention can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, *e.g.*, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al.* (1997) *Nature* 385:810-813 and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_o phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

The 1465, 1587, 2146, 2207, 32838, 336 OR 52908 transgenic animals that express 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA or a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 peptide (detected immunocytochemically, using antibodies directed against 1465, 1587, 2146, 2207, 32838, 336 OR 52908 epitopes) at easily detectable levels should then be further evaluated to identify those animals which display characteristic pain.

B. Cell-Based Systems

Cells that contain and express 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene sequences which encode a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, and, further, exhibit cellular phenotypes associated with nociception, may be used to identify compounds that exhibit analgesic effect. Such cells may include non-recombinant monocyte cell lines, such as U937 (ATCC# CRL-1593), THP-1 (ATCC#TIB-202), and P388D1 (ATCC# TIB-63); endothelial cells such as human umbilical vein endothelial cells (HUVECs), human microvascular endothelial cells (HMVEC), and bovine aortic endothelial cells (BAECs); as well as generic mammalian cell lines such as HeLa cells and COS cells, *e.g.*, COS-7 (ATCC# CRL-1651), and neural cell lines.. Further, such cells may include recombinant, transgenic cell lines. For example, the pain animal models of the invention, discussed above, may be used to generate cell lines, containing one or more cell types involved in nociception, that can be used as cell culture models for this disorder. While primary cultures derived from the pain model transgenic animals of the invention may be utilized, the generation of continuous cell lines is preferred. For examples of techniques which may be used to derive a continuous cell line from the transgenic animals, see Small *et al.*, (1985) *Mol. Cell Biol.* 5:642-648.

Alternatively, cells of a cell type known to be involved in nociception may be transfected with sequences capable of increasing or decreasing the amount of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene expression within the cell. For example, 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene sequences may be introduced into, and overexpressed in, the genome of the cell of interest, or, if endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene sequences are present, they may be either overexpressed or, alternatively disrupted in order to underexpress or inactivate 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene expression.

In order to overexpress a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene, the coding portion of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene may be ligated to a regulatory sequence which is capable of driving gene expression in the cell type of interest, *e.g.*, an endothelial cell. Such regulatory regions will be well known to those of skill in the art, and may be utilized in the absence of undue experimentation. Recombinant methods for expressing target genes are described above.

For underexpression of an endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene sequence, such a sequence may be isolated and engineered such that when reintroduced into the genome of the cell type of interest, the endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 alleles will be inactivated. Preferably, the engineered 1465, 1587, 2146, 2207, 32838, 336 OR 52908 sequence is introduced via gene targeting such that the endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 sequence is disrupted upon integration of the engineered 1465, 1587, 2146, 2207, 32838, 336 OR 52908 sequence into the cell's genome. Transfection of host cells with 1465, 1587, 2146, 2207, 32838, 336 OR 52908 genes is discussed, above.

Cells treated with compounds or transfected with 1465, 1587, 2146, 2207, 32838, 336 OR 52908 genes can be examined for phenotypes associated with nociception.

Transfection of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid may be accomplished by using standard techniques (described in, for example, Ausubel (1989) *supra*). Transfected cells should be evaluated for the presence of the recombinant 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene sequences, for expression and accumulation of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA, and for the presence of recombinant 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein production. In instances wherein a decrease in 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene expression is desired, standard techniques may be used to demonstrate whether a decrease

in endogenous 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene expression and/or in 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein production is achieved.

III. Predictive Medicine:

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein and/or nucleic acid expression as well as 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity, in the context of a biological sample (*e.g.*, blood, serum, cells, *e.g.*, endothelial cells, or tissue, *e.g.*, vascular tissue) to thereby determine whether an individual is afflicted with a predisposition or is experiencing pain. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a painful disorder. For example, mutations in a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene can be assayed for in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby phophylactically treat an individual prior to the onset of a painful disorder.

Another aspect of the invention pertains to monitoring the influence of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulators (*e.g.*, anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibodies or 1465, 1587, 2146, 2207, 32838, 336 OR 52908 ribozymes) on the expression or activity of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 in clinical trials.

These and other agents are described in further detail in the following sections.

A. Diagnostic Assays

To determine whether a subject is afflicted with a disease, a biological sample may be obtained from a subject and the biological sample may be contacted with a compound or an agent capable of detecting a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or nucleic acid (*e.g.*, mRNA or genomic DNA) that encodes a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, in the biological sample. A preferred agent for detecting 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA or genomic DNA. The nucleic acid probe can be, for example, the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid set forth in SEQ ID NO:1, 4, 7, 10,

13, 16 or 19, or a portion thereof, such as an oligonucleotide of at least 15, 20, 25, 30, 25, 40, 45, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein in a sample is an antibody capable of binding to 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')2) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin.

The term "biological sample" is intended to include tissues, cells, and biological fluids isolated from a subject, as well as tissues, cells, and fluids present within a subject. That is, the detection method of the invention can be used to detect 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein include introducing into a subject a labeled anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, mRNA, or

genomic DNA, such that the presence of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, mRNA or genomic DNA in the control sample with the presence of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, mRNA or genomic DNA in the test sample.

B. Prognostic Assays

The present invention further pertains to methods for identifying subjects having or at risk of developing a disease associated with aberrant 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression or activity.

As used herein, the term "aberrant" includes a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression or activity which deviates from the wild type 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression or activity. Aberrant expression or activity includes increased or decreased expression or activity, as well as expression or activity which does not follow the wild type developmental pattern of expression or the subcellular pattern of expression. For example, aberrant 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression or activity is intended to include the cases in which a mutation in the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene causes the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene to be under-expressed or over-expressed and situations in which such mutations result in a non-functional 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or a protein which does not function in a wild-type fashion, *e.g.*, a protein which does not interact with a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 substrate, or one which interacts with a non-1465, 1587, 2146, 2207, 32838, 336 OR 52908 substrate.

The assays described herein, such as the preceding diagnostic assays or the following assays, can be used to identify a subject having or at risk of developing a disease. A biological sample may be obtained from a subject and tested for the presence or absence of a genetic alteration. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene, 2) an addition of one or more nucleotides to a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene, 3) a substitution of one or more nucleotides of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene, 4) a chromosomal rearrangement of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene, 5) an alteration in the level of a messenger RNA transcript of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene, 6) aberrant modification of a 1465, 1587, 2146, 2207, 32838,

336 OR 52908 gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene, 8) a non-wild type level of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -protein, 9) allelic loss of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene, and 10) inappropriate post-translational modification of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -protein.

As described herein, there are a large number of assays known in the art which can be used for detecting genetic alterations in a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene. For example, a genetic alteration in a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene may be detected using a probe/primer in a polymerase chain reaction (PCR) (see, *e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, *e.g.*, Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene (see Abravaya *et al.* (1995) *Nucleic Acids Res.* 23:675-682). This method includes collecting a biological sample from a subject, isolating nucleic acid (*e.g.*, genomic DNA, mRNA or both) from the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene under conditions such that hybridization and amplification of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, D.Y. *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. *et al.* (1988) *Bio-Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene from a biological sample can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in 1465, 1587, 2146, 2207, 32838, 336 OR 52908 can be identified by hybridizing biological sample derived and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotide probes (Cronin, M.T. *et al.* (1996) *Human Mutation* 7:244-255; Kozal, M.J. *et al.* (1996) *Nature Medicine* 2:753-759). For example, genetic mutations in 1465, 1587, 2146, 2207, 32838, 336 OR 52908 can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, M.T. *et al.* (1996) *supra.* Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential, overlapping probes. This step allows for the identification of point mutations. This step is followed by a second hybridization array that allows for the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene in a biological sample and detect mutations by comparing the sequence of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 in the biological sample with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert (1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger (1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays (Naeve, C. W. (1995) *Biotechniques* 19:448-53), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publication No. WO 94/16101; Cohen

et al. (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes formed by hybridizing (labeled) RNA or DNA containing the wild-type 1465, 1587, 2146, 2207, 32838, 336 OR 52908 sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digest the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al.* (1988) *Proc. Natl Acad Sci USA* 85:4397 and Saleeba *et al.* (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in 1465, 1587, 2146, 2207, 32838, 336 OR 52908 cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 sequence, *e.g.*, a wild-type 1465, 1587, 2146, 2207, 32838, 336 OR 52908 sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in 1465, 1587, 2146, 2207, 32838, 336 OR 52908 genes. For example,

single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl. Acad. Sci USA*: 86:2766; see also Cotton (1993) *Mutat. Res.* 285:125-144 and Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-stranded DNA fragments of sample and control 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet* 7:5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to ensure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et al.* (1989) *Proc. Natl Acad. Sci USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential

hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, or small molecule) to effectively treat a disease.

C. Monitoring of Effects During Clinical Trials

The present invention further provides methods for determining the effectiveness of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator (*e.g.*, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator identified herein) in treating a disease. For example, the effectiveness of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator in increasing 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene expression, protein levels, or in upregulating 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity, can be monitored in clinical trials of subjects exhibiting decreased 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene expression, protein levels, or downregulated 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity. Alternatively, the effectiveness of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator in decreasing 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene expression, protein levels, or in downregulating 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity, can be monitored in clinical trials of subjects exhibiting increased 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene expression, protein levels, or 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity. In such clinical trials, the expression or activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene, and preferably, other genes that have been implicated in nociception can be used as a "read out" or marker of the phenotype of a particular cell.

For example, and not by way of limitation, genes, including 1465, 1587, 2146, 2207, 32838, 336 OR 52908 , that are modulated in cells by treatment with an agent which modulates 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents which modulate 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity on subjects suffering from a painful disorder in, for example, a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 and other genes implicated in the painful disorder. The levels of gene expression (*e.g.*, a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods described herein, or by measuring the levels of activity of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent which modulates 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity. This response state may be determined before, and at various points during treatment of the individual with the agent which modulates 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent which modulates 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, or small molecule identified by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, mRNA, or genomic DNA in the pre-administration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, mRNA, or genomic DNA in the pre-administration sample with the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 to higher levels than detected, *i.e.*, to increase the effectiveness of the agent.

Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 to lower levels than detected, *i.e.* to decrease the effectiveness of the agent. According to such an embodiment, 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression or activity may be used as an indicator of the effectiveness of an agent, even in the absence of an observable phenotypic response.

IV. Methods of Treatment:

The present invention provides for both prophylactic and therapeutic methods of treating a subject, *e.g.*, a human, at risk of (or susceptible to) a disease. With regard to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics," as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers to the study of how a patient's genes determine his or her response to a drug (*e.g.*, a patient's "drug response phenotype", or "drug response genotype").

Thus, another aspect of the invention provides methods for tailoring an subject's prophylactic or therapeutic treatment with either the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 molecules of the present invention or 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

A. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease by administering to the subject an agent which modulates 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression or 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity. Subjects at risk for a cardiovascular disease, *e.g.*, atherosclerosis and/or thrombosis, can be identified by, for example, any or a combination of the diagnostic or prognostic assays described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of aberrant 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression or activity, such that a disease is prevented or, alternatively, delayed

in its progression. Depending on the type of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 aberrancy, for example, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 , 1465, 1587, 2146, 2207, 32838, 336 OR 52908 agonist or 1465, 1587, 2146, 2207, 32838, 336 OR 52908 antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

B. Therapeutic Methods

Described herein are methods and compositions whereby pain may be ameliorated. Certain painful disorders are brought about, at least in part, by an excessive level of a gene product, or by the presence of a gene product exhibiting an abnormal or excessive activity. As such, the reduction in the level and/or activity of such gene products would bring about the amelioration of pain. Techniques for the reduction of gene expression levels or the activity of a protein are discussed below.

Alternatively, certain other painful disorders are brought about, at least in part, by the absence or reduction of the level of gene expression, or a reduction in the level of a protein's activity. As such, an increase in the level of gene expression and/or the activity of such proteins would bring about the amelioration of pain.

In some cases, the up-regulation of a gene in a disease state reflects a protective role for that gene product in responding to the disease condition. Enhancement of such a gene's expression, or the activity of the gene product, will reinforce the protective effect it exerts. Some pain states may result from an abnormally low level of activity of such a protective gene. In these cases also, an increase in the level of gene expression and/or the activity of such gene products would bring about the amelioration of pain. Techniques for increasing target gene expression levels or target gene product activity levels are discussed herein.

Accordingly, another aspect of the invention pertains to methods of modulating 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell with a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 or agent that modulates one or more of the activities of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein activity associated with the cell (*e.g.*, an endothelial cell or an ovarian cell). An agent that modulates 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring target molecule of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908

protein (*e.g.*, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 ligand or substrate), a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibody, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 agonist or antagonist, a peptidomimetic of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 agonist or antagonist, or other small molecule. In one embodiment, the agent stimulates one or more 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activities. Examples of such stimulatory agents include active 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein and a nucleic acid molecule encoding 1465, 1587, 2146, 2207, 32838, 336 OR 52908 that has been introduced into the cell. In another embodiment, the agent inhibits one or more 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activities. Examples of such inhibitory agents include antisense 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid molecules, anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibodies, and 1465, 1587, 2146, 2207, 32838, 336 OR 52908 inhibitors. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant or unwanted expression or activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression or activity. In another embodiment, the method involves administering a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression or activity.

Stimulation of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity is desirable in situations in which 1465, 1587, 2146, 2207, 32838, 336 OR 52908 is abnormally downregulated and/or in which increased 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity is likely to have a beneficial effect. Likewise, inhibition of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity is desirable in situations in which 1465, 1587, 2146, 2207, 32838, 336 OR 52908 is abnormally upregulated and/or in which decreased 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity is likely to have a beneficial effect.

(i) Methods for Inhibiting Target Gene Expression, Synthesis, or Activity

As discussed above, genes involved in cardiovascular disorders may cause such disorders via an increased level of gene activity. In some cases, such up-regulation may

have a causative or exacerbating effect on the disease state. A variety of techniques may be used to inhibit the expression, synthesis, or activity of such genes and/or proteins.

For example, compounds such as those identified through assays described above, which exhibit inhibitory activity, may be used in accordance with the invention to ameliorate pain. Such molecules may include, but are not limited to, small organic molecules, peptides, antibodies, and the like.

For example, compounds can be administered that compete with endogenous ligand for the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. The resulting reduction in the amount of ligand-bound 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein will modulate endothelial cell physiology. Compounds that can be particularly useful for this purpose include, for example, soluble proteins or peptides, such as peptides comprising one or more of the extracellular domains, or portions and/or analogs thereof, of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, including, for example, soluble fusion proteins such as Ig-tailed fusion proteins. (For a discussion of the production of Ig-tailed fusion proteins, see, for example, U.S. Pat. No. 5,116,964). Alternatively, compounds, such as ligand analogs or antibodies, that bind to the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 receptor site, but do not activate the protein, (*e.g.*, receptor-ligand antagonists) can be effective in inhibiting 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein activity.

Further, antisense and ribozyme molecules which inhibit expression of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene may also be used in accordance with the invention to inhibit aberrant 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene activity. Still further, triple helix molecules may be utilized in inhibiting aberrant 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene activity.

The antisense nucleic acid molecules used in the methods of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and

then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, an antisense nucleic acid molecule used in the methods of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids. Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

In still another embodiment, an antisense nucleic acid used in the methods of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA transcripts to thereby inhibit translation of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA. A ribozyme having specificity for a 1465, 1587, 2146, 2207, 32838, 336 OR 52908-encoding nucleic acid can be designed based upon the nucleotide sequence of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 cDNA disclosed herein (*i.e.*, SEQ ID NO:1 or 3). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a 1465, 1587, 2146, 2207, 32838, 336 OR 52908-encoding mRNA (see, for example, Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742). Alternatively, 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules (see, for example, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418).

1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene expression can also be inhibited by targeting nucleotide sequences complementary to the regulatory region of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 (*e.g.*, the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 promoter and/or enhancers) to form triple helical structures that prevent transcription of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene in target cells (see, for example, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15).

Antibodies that are both specific for the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein and interfere with its activity may also be used to modulate or inhibit 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein function. Such antibodies may be generated using standard techniques described herein, against the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein itself or against peptides corresponding to portions of the protein. Such antibodies include but are not limited to polyclonal, monoclonal, Fab fragments, single chain antibodies, or chimeric antibodies.

In instances where the target gene protein is intracellular and whole antibodies are used, internalizing antibodies may be preferred. Lipofectin liposomes may be used to deliver the antibody or a fragment of the Fab region which binds to the target epitope into cells. Where fragments of the antibody are used, the smallest inhibitory fragment which binds to the target protein's binding domain is preferred. For example, peptides having an amino acid sequence corresponding to the domain of the variable region of the antibody that binds to the target gene protein may be used. Such peptides may be synthesized chemically or produced via recombinant DNA technology using methods well known in the art (described in, for example, Creighton (1983), *supra*; and Sambrook *et al.* (1989) *supra*). Single chain neutralizing antibodies which bind to intracellular target gene epitopes may also be administered. Such single chain antibodies may be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population by utilizing, for example, techniques such as those described in Marasco *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:7889-7893).

In some instances, the target gene protein is extracellular, or is a transmembrane protein, such as the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. Antibodies that are specific for one or more extracellular domains of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, for example, and that interfere with its activity, are particularly useful in treating pain or a painful disorder. Such antibodies are especially

efficient because they can access the target domains directly from the bloodstream. Any of the administration techniques described below which are appropriate for peptide administration may be utilized to effectively administer inhibitory target gene antibodies to their site of action.

(ii) Methods for Restoring or Enhancing Target Gene Activity

Genes that cause pain may be underexpressed within cardiovascular disease situations. Alternatively, the activity of the protein products of such genes may be decreased, leading to the development of pain. Such down-regulation of gene expression or decrease of protein activity might have a causative or exacerbating effect on the disease state.

In some cases, genes that are up-regulated in the disease state might be exerting a protective effect. A variety of techniques may be used to increase the expression, synthesis, or activity of genes and/or proteins that exert a protective effect in response to pain conditions.

Described in this section are methods whereby the level 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity may be increased to levels wherein pain are ameliorated. The level of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity may be increased, for example, by either increasing the level of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene expression or by increasing the level of active 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein which is present.

For example, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, at a level sufficient to ameliorate pain may be administered to a patient exhibiting such symptoms. Any of the techniques discussed below may be used for such administration. One of skill in the art will readily know how to determine the concentration of effective, non-toxic doses of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, utilizing techniques such as those described below.

Additionally, RNA sequences encoding a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein may be directly administered to a patient exhibiting pain, at a concentration sufficient to produce a level of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein such that pain are ameliorated. Any of the techniques discussed below, which achieve intracellular administration of compounds, such as, for example, liposome administration, may be used for the administration of such RNA molecules. The RNA molecules may be produced, for example, by recombinant techniques such as those described herein.

Further, subjects may be treated by gene replacement therapy. One or more copies of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene, or a portion thereof, that directs the production of a normal 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein with 1465, 1587, 2146, 2207, 32838, 336 OR 52908 function, may be inserted into cells using vectors which include, but are not limited to adenovirus, adeno-associated virus, and retrovirus vectors, in addition to other particles that introduce DNA into cells, such as liposomes. Additionally, techniques such as those described above may be used for the introduction of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene sequences into human cells.

Cells, preferably, autologous cells, containing 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expressing gene sequences may then be introduced or reintroduced into the subject at positions which allow for the amelioration of pain. Such cell replacement techniques may be preferred, for example, when the gene product is a secreted, extracellular gene product.

C. Pharmaceutical Compositions

Another aspect of the invention pertains to methods for treating a subject suffering from a disease. These methods involve administering to a subject an agent which modulates 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression or activity (*e.g.*, an agent identified by a screening assay described herein), or a combination of such agents. In another embodiment, the method involves administering to a subject a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted 1465, 1587, 2146, 2207, 32838, 336 OR 52908 expression or activity.

Stimulation of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity is desirable in situations in which 1465, 1587, 2146, 2207, 32838, 336 OR 52908 is abnormally downregulated and/or in which increased 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity is likely to have a beneficial effect. Likewise, inhibition of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity is desirable in situations in which 1465, 1587, 2146, 2207, 32838, 336 OR 52908 is abnormally upregulated and/or in which decreased 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity is likely to have a beneficial effect.

The agents which modulate 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity can be administered to a subject using pharmaceutical compositions suitable for such administration. Such compositions typically comprise the agent (*e.g.*, nucleic acid

molecule, protein, or antibody) and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition used in the therapeutic methods of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyetheylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can

be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, and sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the agent that modulates 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity (*e.g.*, a fragment of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or an anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The agents that modulate 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the agents that modulate 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the agent that modulates 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an agent for the treatment of subjects.

Toxicity and therapeutic efficacy of such agents can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining

the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and can be expressed as the ratio LD50/ED50. Agents which exhibit large therapeutic indices are preferred. While agents that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such agents to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulating agents lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any agent used in the therapeutic methods of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

As defined herein, a therapeutically effective amount of protein or polypeptide (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, preferably about 0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and even more preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The skilled artisan will appreciate that certain factors may influence the dosage required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

In a preferred example, a subject is treated with antibody, protein, or polypeptide in the range of between about 0.1 to 20 mg/kg body weight, one time per week for between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. It will also be appreciated that the effective dosage of antibody, protein, or polypeptide used for treatment may

increase or decrease over the course of a particular treatment. Changes in dosage may result and become apparent from the results of diagnostic assays as described herein.

The present invention encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (*i.e*,. including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds. It is understood that appropriate doses of small molecule agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of the small molecule will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention.

Exemplary doses include milligram or microgram amounts of the small molecule per kilogram of subject or sample weight (*e.g.*, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram). It is furthermore understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. Such appropriate doses may be determined using the assays described herein. When one or more of these small molecules is to be administered to an animal (*e.g.*, a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of

administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

Further, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*, methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (*e.g.*, dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophase colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon *et al.*, "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld *et al.* (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom *et al.*, "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson *et al.* (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A

Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera *et al.* (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin *et al.* (eds.), pp. 303-16 (Academic Press 1985), and Thorpe *et al.*, "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev., 62:119-58 (1982). Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

The nucleic acid molecules used in the methods of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent 5,328,470) or by stereotactic injection (see, *e.g.*, Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

D. Pharmacogenomics

In conjunction with the therapeutic methods of the invention, pharmacogenomics (*i.e.*, the study of the relationship between a subject's genotype and that subject's response to a foreign compound or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer an agent which modulates 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity, as well as tailoring the dosage and/or therapeutic regimen of treatment with an agent which modulates 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum, M. *et al.* (1996) *Clin. Exp.Pharmacol. Physiol.* 23(10-11): 983-985 and Linder, M.W. *et al.* (1997) *Clin. Chem.* 43(2):254-266. In general, two types

of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate aminopeptidase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

One pharmacogenomics approach to identifying genes that predict drug response, known as "a genome-wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (*e.g.*, a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which has two variants). Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a high resolution map can be generated from a combination of some ten million known single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast majority may not be disease-associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

Alternatively, a method termed the "candidate gene approach" can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drug target is known (*e.g.*, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein used in the methods of the present invention), all common variants of that gene can be fairly easily identified in the population and it can be determined if having one version of the gene versus another is associated with a particular drug response.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2)

and the cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Alternatively, a method termed the "gene expression profiling" can be utilized to identify genes that predict drug response. For example, the gene expression of an animal dosed with a drug (*e.g.*, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 molecule or 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator used in the methods of the present invention) can give an indication whether gene pathways related to toxicity have been turned on.

Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment of a subject. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and, thus, enhance therapeutic or prophylactic efficiency when treating a subject suffering from a cardiovascular disease, *e.g.*, atherosclerosis, with an agent which modulates 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity.

V. Recombinant Expression Vectors and Host Cells Used in the Methods of the Invention

The methods of the invention (*e.g.*, the screening assays described herein) include the use of vectors, preferably expression vectors, containing a nucleic acid encoding a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another

nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors to be used in the methods of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel (1990) *Methods Enzymol.* 185:3-7. Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce

proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins, mutant forms of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins, fusion proteins, and the like).

The recombinant expression vectors to be used in the methods of the invention can be designed for expression of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins in prokaryotic or eukaryotic cells. For example, 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors), yeast cells, or mammalian cells. Suitable host cells are discussed further in Goeddel (1990) *supra*. Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Purified fusion proteins can be utilized in 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity assays, (*e.g.*, direct assays or competitive assays described in detail below), or to generate antibodies specific for 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins. In a preferred embodiment, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 fusion protein expressed in a retroviral expression vector of the present invention can be utilized to infect bone marrow cells which are subsequently transplanted into irradiated

recipients. The pathology of the subject recipient is then examined after sufficient time has passed (*e.g.*, six weeks).

In another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J. *et al.*, *Molecular Cloning: A Laboratory Manual. 2nd ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid).

The methods of the invention may further use a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific, or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid, or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes, see Weintraub, H. *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to the use of host cells into which a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid molecule of the invention is introduced, *e.g.*, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid molecule

within a recombinant expression vector or a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid molecule containing sequences which allow it to homologously recombine into a specific site of the host cell's genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook *et al.* (*Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

A host cell used in the methods of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. Accordingly, the invention further provides methods for producing a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention (into which a recombinant expression vector encoding a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein has been introduced) in a suitable medium such that a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein is produced. In another embodiment, the method further comprises isolating a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein from the medium or the host cell.

VI. Isolated Nucleic Acid Molecules Used in the Methods of the Invention

The methods of the invention include the use of isolated nucleic acid molecules that encode 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes to identify 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -encoding nucleic acid molecules (e.g., 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA) and fragments for use as PCR primers for the amplification or mutation of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

A nucleic acid molecule used in the methods of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1, 4, or 7 , or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or portion of the nucleic acid sequence of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, as a hybridization probe, 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19 can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19.

A nucleic acid used in the methods of the invention can be amplified using cDNA, mRNA or, alternatively, genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. Furthermore, oligonucleotides corresponding to 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In a preferred embodiment, the isolated nucleic acid molecules used in the methods of the invention comprise the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, a complement of the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or

19, or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, is one which is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19 such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19 thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule used in the methods of the present invention comprises a nucleotide sequence which is at least about 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the entire length of the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, or a portion of any of this nucleotide sequence.

Moreover, the nucleic acid molecules used in the methods of the invention can comprise only a portion of the nucleic acid sequence of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, , for example, a fragment which can be used as a probe or primer or a fragment encoding a portion of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, *e.g.*, a biologically active portion of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 consecutive nucleotides of a sense sequence of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, of an anti-sense sequence of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, . In one embodiment, a nucleic acid molecule used in the methods of the present invention comprises a nucleotide sequence which is greater than 100, 100-200, 200-300, 300-400, 400-500, 500-600, 600-700, 700-800, 800-900, 900-1000, 1000-1100, 1100-1200, 1200-1300, or more nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences that are significantly identical or homologous to each other remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85% or 90% identical to each other remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), sections 2, 4 and 6. Additional stringent conditions can

be found in *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), chapters 7, 9 and 11. A preferred, non-limiting example of stringent hybridization conditions includes hybridization in 4X sodium chloride/sodium citrate (SSC), at about 65-70°C (or hybridization in 4X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 1X SSC, at about 65-70° C. A preferred, non-limiting example of highly stringent hybridization conditions includes hybridization in 1X SSC, at about 65-70°C (or hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 0.3X SSC, at about 65-70°C. A preferred, non-limiting example of reduced stringency hybridization conditions includes hybridization in 4X SSC, at about 50-60°C (or alternatively hybridization in 6X SSC plus 50% formamide at about 40-45°C) followed by one or more washes in 2X SSC, at about 50-60°C. Ranges intermediate to the above-recited values, *e.g.*, at 65-70°C or at 42-50°C are also intended to be encompassed by the present invention. SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes each after hybridization is complete. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, $T_m(°C) = 2(\#\text{ of A + T bases}) + 4(\#\text{ of G + C bases})$. For hybrids between 18 and 49 base pairs in length, $T_m(°C) = 81.5 + 16.6(\log_{10}[Na^+]) + 0.41(\%G+C) - (600/N)$, where N is the number of bases in the hybrid, and $[Na^+]$ is the concentration of sodium ions in the hybridization buffer ($[Na^+]$ for 1xSSC = 0.165 M). It will also be recognized by the skilled practitioner that additional reagents may be added to hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (*e.g.*, BSA or salmon or herring sperm carrier DNA), detergents (*e.g.*, SDS), chelating agents (*e.g.*, EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional preferred, non-limiting example of stringent hybridization conditions is hybridization in 0.25-0.5M NaH_2PO_4, 7% SDS at about 65°C, followed by one or more washes at 0.02M NaH_2PO_4, 1% SDS at 65°C, see *e.g.*, Church and Gilbert (1984) *Proc. Natl. Acad. Sci. USA* 81:1991-1995, (or alternatively 0.2X SSC, 1% SDS).

In preferred embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, such as by measuring a level of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA levels or determining whether a genomic 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene has been mutated or deleted.

The methods of the invention further encompass the use of nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, due to degeneracy of the genetic code and thus encode the same 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins as those encoded by the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, . In another embodiment, an isolated nucleic acid molecule included in the methods of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO:3, 6, 9, 12, 15, 18 or 21 .

The methods of the invention further include the use of allelic variants of human 1465, 1587, 2146, 2207, 32838, 336 OR 52908 , *e.g.*, functional and non-functional allelic variants. Functional allelic variants are naturally occurring amino acid sequence variants of the human 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein that maintain a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:3, 6, 9, 12, 15, 18 or 21 , or substitution, deletion or insertion of non-critical residues in non-critical regions of the protein.

Non-functional allelic variants are naturally occurring amino acid sequence variants of the human 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein that do not have a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity. Non-functional allelic variants will typically contain a non-conservative substitution, deletion, or insertion or premature truncation of the amino acid sequence of SEQ ID NO:3, 6, 9, 12, 15, 18 or 21 , or a substitution, insertion or deletion in critical residues or critical regions of the protein.

The methods of the present invention may further use non-human orthologues of the human 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. Orthologues of the human 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein are proteins that are isolated from non-human organisms and possess the same 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity.

The methods of the present invention further include the use of nucleic acid molecules comprising the nucleotide sequence of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, or a portion thereof, in which a mutation has been introduced. The mutation may lead to amino acid substitutions at "non-essential" amino acid residues or at "essential" amino acid residues. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 (*e.g.*, the sequence of SEQ ID NO:3, 6, 9, 12, 15, 18 or 21) without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins of the present invention are not likely to be amenable to alteration.

Mutations can be introduced into SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, glycine, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for 1465, 1587, 2146, 2207, 32838, 336 OR 52908 biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, the encoded protein can be expressed recombinantly and the activity of the protein can be determined using the assay described herein.

Another aspect of the invention pertains to the use of isolated nucleic acid molecules which are antisense to the nucleotide sequence of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, . An "antisense" nucleic acid comprises a nucleotide sequence which is

complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire 1465, 1587, 2146, 2207, 32838, 336 OR 52908 coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 . The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding 1465, 1587, 2146, 2207, 32838, 336 OR 52908 . The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding 1465, 1587, 2146, 2207, 32838, 336 OR 52908 disclosed herein, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-

methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest). Antisense nucleic acid molecules used in the methods of the invention are further described above, in section IV.

In yet another embodiment, the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid molecules used in the methods of the present invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup B. *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4 (1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *et al.* (1996) *Proc. Natl. Acad. Sci.* 93:14670-675.

PNAs of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid molecules can be used in the therapeutic and diagnostic applications described herein. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (*e.g.*, by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (*e.g.*, S1 nucleases (Hyrup B. *et al.* (1996) *supra*)); or as probes or

primers for DNA sequencing or hybridization (Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *et al.* (1996) *supra*).

In another embodiment, PNAs of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 can be modified, (*e.g.*, to enhance their stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid molecules can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (*e.g.*, RNAse H and DNA polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup B. *et al.* (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup B. *et al.* (1996) *supra* and Finn P.J. *et al.* (1996) *Nucleic Acids Res.* 24 (17): 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag, M. *et al.* (1989) *Nucleic Acid Res.* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn P.J. *et al.* (1996) *supra*). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser, K.H. *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5: 1119-11124).

In other embodiments, the oligonucleotide used in the methods of the invention may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, *e.g.*, Krol *et al.* (1988) *Bio-Techniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, (*e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

VII. Isolated 1465, 1587, 2146, 2207, 32838, 336 OR 52908 Proteins and Anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 Antibodies Used in the Methods of the Invention

The methods of the invention include the use of isolated 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibodies. In one embodiment, native 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

As used herein, a "biologically active portion" of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein includes a fragment of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein having a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity. Biologically active portions of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein include peptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:3, 6, 9, 12, 15, 18 or 21 , which include fewer amino acids than the full length 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins, and exhibit at least one activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein (*e.g.*, the N-terminal region of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein that is believed to be involved in the regulation of apoptotic activity). A biologically active portion of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein can be a polypeptide which is, for example, 25, 50, 75, 100, 125, 150, 175, 200, 250, 300 or more amino acids in length. Biologically active portions of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein can be used as targets for developing agents which modulate a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 activity.

In a preferred embodiment, the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein used in the methods of the invention has an amino acid sequence shown in SEQ ID NO:3, 6, 9, 12, 15, 18 or 21 . In other embodiments, the 1465, 1587, 2146, 2207, 32838,

336 OR 52908 protein is substantially identical to SEQ ID NO:3, 6, 9, 12, 15, 18 or 21 , and retains the functional activity of the protein of SEQ ID NO:3, 6, 9, 12, 15, 18 or 21 , yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail in subsection V above. Accordingly, in another embodiment, the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein used in the methods of the invention is a protein which comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NO:3, 6, 9, 12, 15, 18 or 21 .

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-identical sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence (*e.g.*, when aligning a second sequence to the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 amino acid sequence of SEQ ID NO:3, 6, 9, 12, 15, 18 or 21 having 500 amino acid residues, at least 75, preferably at least 150, more preferably at least 225, even more preferably at least 300, and even more preferably at least 400 or more amino acid residues are aligned). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* 48:444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package, using either a Blosum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length

weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package, using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (*Comput. Appl. Biosci.* 4:11-17 (1988)) which has been incorporated into the ALIGN program (version 2.0 or 2.0U), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The methods of the invention may also use 1465, 1587, 2146, 2207, 32838, 336 OR 52908 chimeric or fusion proteins. As used herein, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 "chimeric protein" or "fusion protein" comprises a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide operatively linked to a non-1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide. An "1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 molecule, whereas a "non-1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, *e.g.*, a protein which is different from the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein and which is derived from the same or a different organism. Within a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 fusion protein the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide can correspond to all or a portion of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. In a preferred embodiment, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 fusion protein comprises at least one biologically active portion of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. In another preferred embodiment, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 fusion protein comprises at least two biologically active portions of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide and the non-1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide are fused in-frame to each other. The non-1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide can be fused to the N-terminus or C-terminus of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide.

For example, in one embodiment, the fusion protein is a GST-1465, 1587, 2146, 2207, 32838, 336 OR 52908 fusion protein in which the 1465, 1587, 2146, 2207, 32838,

336 OR 52908 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant 1465, 1587, 2146, 2207, 32838, 336 OR 52908 .

In another embodiment, this fusion protein is a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 can be increased through use of a heterologous signal sequence.

The 1465, 1587, 2146, 2207, 32838, 336 OR 52908 fusion proteins used in the methods of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The 1465, 1587, 2146, 2207, 32838, 336 OR 52908 fusion proteins can be used to affect the bioavailability of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 substrate. Use of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 fusion proteins may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein; (ii) mis-regulation of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 gene; and (iii) aberrant post-translational modification of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein.

Moreover, the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -fusion proteins used in the methods of the invention can be used as immunogens to produce anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibodies in a subject, to purify 1465, 1587, 2146, 2207, 32838, 336 OR 52908 ligands and in screening assays to identify molecules which inhibit the interaction of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 with a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 substrate.

Preferably, a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 chimeric or fusion protein used in the methods of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between

two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein.

The present invention also pertains to the use of variants of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins which function as either 1465, 1587, 2146, 2207, 32838, 336 OR 52908 agonists (mimetics) or as 1465, 1587, 2146, 2207, 32838, 336 OR 52908 antagonists. Variants of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. An agonist of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. An antagonist of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein can inhibit one or more of the activities of the naturally occurring form of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein by, for example, competitively modulating a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 -mediated activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein.

In one embodiment, variants of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein which function as either 1465, 1587, 2146, 2207, 32838, 336 OR 52908 agonists (mimetics) or as 1465, 1587, 2146, 2207, 32838, 336 OR 52908 antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein for 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein agonist or antagonist activity. In one embodiment, a variegated library of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 variants can be produced by, for example, enzymatically ligating a mixture of synthetic

oligonucleotides into gene sequences such that a degenerate set of potential 1465, 1587, 2146, 2207, 32838, 336 OR 52908 sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 sequences therein. There are a variety of methods which can be used to produce libraries of potential 1465, 1587, 2146, 2207, 32838, 336 OR 52908 variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential 1465, 1587, 2146, 2207, 32838, 336 OR 52908 sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477).

In addition, libraries of fragments of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein coding sequence can be used to generate a variegated population of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 fragments for screening and subsequent selection of variants of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 proteins. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming

appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify 1465, 1587, 2146, 2207, 32838, 336 OR 52908 variants (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave *et al.* (1993) *Protein Engineering* 6(3):327-331).

The methods of the present invention further include the use of anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibodies. An isolated 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind 1465, 1587, 2146, 2207, 32838, 336 OR 52908 using standard techniques for polyclonal and monoclonal antibody preparation. A full-length 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein can be used or, alternatively, antigenic peptide fragments of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 can be used as immunogens. The antigenic peptide of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:3, 6, 9, 12, 15, 18 or 21 and encompasses an epitope of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 such that an antibody raised against the peptide forms a specific immune complex with the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

Preferred epitopes encompassed by the antigenic peptide are regions of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 that are located on the surface of the protein, *e.g.*, hydrophilic regions, as well as regions with high antigenicity.

A 1465, 1587, 2146, 2207, 32838, 336 OR 52908 immunogen is typically used to prepare antibodies by immunizing a suitable subject, (*e.g.*, rabbit, goat, mouse, or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein or a chemically synthesized 1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic 1465, 1587, 2146, 2207, 32838, 336 OR 52908 preparation induces a polyclonal anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibody response.

The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 . Examples of immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind 1465, 1587, 2146, 2207, 32838, 336 OR 52908 molecules. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of 1465, 1587, 2146, 2207, 32838, 336 OR 52908 . A monoclonal antibody composition thus typically displays a single binding affinity for a particular 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein with which it immunoreacts.

Polyclonal anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibodies can be prepared as described above by immunizing a suitable subject with a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 immunogen. The anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized 1465, 1587, 2146, 2207, 32838, 336 OR 52908 . If desired, the antibody molecules directed against 1465, 1587, 2146, 2207, 32838, 336 OR 52908 can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, *e.g.*, when the anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497) (see also, Brown *et al.* (1981) *J. Immunol.* 127:539-46; Brown *et al.* (1980) *J. Biol. Chem.* 255:4980-83; Yeh *et al.* (1976) *Proc. Natl. Acad. Sci. USA* 76:2927-31; and Yeh *et al.* (1982) *Int. J. Cancer* 29:269-75), the more recent human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985) *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known (see generally Kenneth, R. H. in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); Lerner, E.

A. (1981) *Yale J. Biol. Med.* 54:387-402; Gefter, M. L. *et al.* (1977) *Somatic Cell Genet.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds 1465, 1587, 2146, 2207, 32838, 336 OR 52908 .

Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 monoclonal antibody (see, *e.g.*, G. Galfre *et al.* (1977) *Nature* 266:55052; Gefter *et al.* (1977) *supra*; Lerner (1981) *supra*; and Kenneth (1980) *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (*e.g.*, a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, *e.g.*, the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind 1465, 1587, 2146, 2207, 32838, 336 OR 52908 , *e.g.*, using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with 1465, 1587, 2146, 2207, 32838, 336 OR 52908 to thereby isolate immunoglobulin library members that bind 1465, 1587, 2146, 2207, 32838, 336 OR 52908 . Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and

the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* PCT International Publication No. WO 92/18619; Dower *et al.* PCT International Publication No. WO 91/17271; Winter *et al.* PCT International Publication WO 92/20791; Markland *et al.* PCT International Publication No. WO 92/15679; Breitling *et al.* PCT International Publication WO 93/01288; McCafferty *et al.* PCT International Publication No. WO 92/01047; Garrard *et al.* PCT International Publication No. WO 92/09690; Ladner *et al.* PCT International Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J. Mol. Biol.* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3576-3580; Garrad *et al.* (1991) *Bio/Technology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc. Acid Res.* 19:4133-4137; Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; and McCafferty *et al.* (1990) *Nature* 348:552-554.

Additionally, recombinant anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the methods of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson *et al.* International Application No. PCT/US86/02269; Akira, *et al.* European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison *et al.* European Patent Application 173,494; Neuberger *et al.* PCT International Publication No. WO 86/01533; Cabilly *et al.* U.S. Patent No. 4,816,567; Cabilly *et al.* European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559; Morrison, S. L. (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; Winter U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

An anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibody can be used to detect 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 protein. Anti-1465, 1587, 2146, 2207, 32838, 336 OR 52908 antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application, as well as the Figure and the Sequence Listing is incorporated herein by reference.

EXAMPLES

EXAMPLE 1: TISSUE DISTRIBUTION OF USING TAQMAN® ANALYSIS

This example describes the TaqMan® procedure. The TaqMan® procedure is a quantitative, reverse transcription PCR-based approach for detecting mRNA. The RT-PCR reaction exploits the 5' nuclease activity of AmpliTaq Gold™ DNA Polymerase to cleave a TaqMan® probe during PCR. Briefly, cDNA was generated from the samples of interest, *e.g.*, heart, kidney, liver, skeletal muscle, and various vessels, and used as the starting material for PCR amplification. In addition to the 5' and 3' gene-specific primers, a gene-specific oligonucleotide probe (complementary to the region being amplified) was included in the reaction (*i.e.*, the TaqMan® probe). The TaqMan® probe includes the

oligonucleotide with a fluorescent reporter dye covalently linked to the 5' end of the probe (such as FAM (6-carboxyfluorescein), TET (6-carboxy-4,7,2',7'-tetrachlorofluorescein), JOE (6-carboxy-4,5-dichloro-2,7-dimethoxyfluorescein), or VIC) and a quencher dye (TAMRA (6-carboxy-N,N,N',N'-tetramethylrhodamine) at the 3' end of the probe.

During the PCR reaction, cleavage of the probe separates the reporter dye and the quencher dye, resulting in increased fluorescence of the reporter. Accumulation of PCR products is detected directly by monitoring the increase in fluorescence of the reporter dye. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence. During PCR, if the target of interest is present, the probe specifically anneals between the forward and reverse primer sites. The 5'-3' nucleolytic activity of the AmpliTaq® Gold DNA Polymerase cleaves the probe between the reporter and the quencher only if the probe hybridizes to the target. The probe fragments are then displaced from the target, and polymerization of the strand continues. The 3' end of the probe is blocked to prevent extension of the probe during PCR. This process occurs in every cycle and does not interfere with the exponential accumulation of product. RNA was prepared using the trizol method and treated with DNase to remove contaminating genomic DNA. cDNA was synthesized using standard techniques. Mock cDNA synthesis in the absence of reverse transcriptase resulted in samples with no detectable PCR amplification of the control gene confirms efficient removal of genomic DNA contamination.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

What is claimed:

1. A method for identifying a compound capable of treating a pain disorder, comprising assaying the ability of the compound to modulate 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid expression or 1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide activity, thereby identifying a compound capable of treating a pain disorder.

2. A method for identifying a compound capable of modulating a pain signaling mechanism comprising:

a) contacting a cell which expresses 1465, 1587, 2146, 2207, 32838, 336 OR 52908 with a test compound; and

b) assaying the ability of the test compound to modulate the expression of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid or the activity of a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide, thereby identifying a compound capable of modulating pain signalling.

3. A method for modulating a pain signaling mechanism in a cell comprising contacting a cell with a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator, thereby modulating a pain signaling mechanism in the cell.

4. The method of claim 2, wherein the cell is a brain cell, neuron, or cell derived from spinal cord or dorsal root ganglion.

5. The method of claim 3, wherein the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator is a small organic molecule, peptide, antibody or antisense nucleic acid molecule.

6. The method of claim 3, wherein the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator is capable of modulating 1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide activity.

7. The method of claim 6, wherein the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator is a small organic molecule, peptide, antibody or antisense nucleic acid molecule.

8. The method of claim 6, wherein the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator is capable of modulating 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid expression.

9. A method for treating a subject having a pain disorder characterized by aberrant 1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide activity or aberrant 1465, 1587, 2146, 2207, 32838, 336 OR 52908 nucleic acid expression comprising administering to the subject a 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator, thereby treating said subject having a pain disorder.

10. The method of claim 9, wherein said pain disorder includes inflammatory pain, chronic pain, neuropathic pain, causalgia, fibromyalgia, cancer pain, migraine/headache pain and tissue pain.

11. The method of claim 9, wherein said 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator is administered in a pharmaceutically acceptable formulation.

12. The method of claim 9, wherein the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator is a small organic molecule, peptide, antibody or antisense nucleic acid molecule.

13. The method of claim 9, wherein the 1465, 1587, 2146, 2207, 32838, 336 OR 52908 modulator is capable of modulating 1465, 1587, 2146, 2207, 32838, 336 OR 52908 polypeptide activity.

SEQUENCE LISTING

<110> Millennium Pharmaceuticals, Inc
Silos-Santiago, Inmaculada
Rosenfeld, Julie Beth

<120> METHODS AND COMPOSITIONS IN TREATING
PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
32838,336 and 52908

<130> MPI01-294P1RNWOM

<150> US 60/341,953
<151> 2001-12-19

<160> 21

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 2113
<212> DNA
<213> Homo Sapiens

<400> 1
gcagccgccg cggtcgggct gccccctccc ctcgccccga ccgctcccct gctggtgagg 60
gcctgcgcag gcggcggcgg cggcgccctt ggtggcagtg gtggcggcgg agcagcctcc 120
cgggatcgtg tctggagctc gaggagaggg tagtgcccgg cgagctatgc acgggggcgg 180
cggcgtctcc tcctgttccg cctcctcagt ctcctcggtc tcgcagaatc cgccggcggc 240
ggcggcgcca ggacatggag ctcgagaaca tcgtggccaa ctcgctgctg ctgaaagcgc 300
gtcaaggagg atatggcaaa aaaagtggtc gtagtaaaaa atggaaggag atactgacac 360
tgcctcctgt cagccagtgc agtgagctta gacattccat tgaaaaggat tatagcagtc 420
tttgtgacaa gcaaccgata ggaagacgtc tcttcaggca gttctgtgat accaaaccca 480
ctctaaagag gcacattgaa ttcttggatg cagtggcaga atatgaagtt gccgatgatg 540
aggaccgaag tgattgtgga ctgtcaatct tagatagatt cttcaatgat aagttggcag 600
ccccttacc agaaatacct ccagatgttg tgacagaatg tagattggga ctgaaggagg 660
agaacccttc caaaaaagcc tttgaggaat gtactagagt tgcccataac tacctaagag 720
gggaaccatt tgaagaatac caagaaagct catatttttc tcagttttta caatggaaat 780
ggctggaaag gcaacccgta acaaagaaca catttagaca ttacagagtt ctaggaaaag 840
gcggatttgg agaggtttgc gcctgtcaag tgcgagccac aggaaaaatg tatgcctgca 900
aaaagctaca aaaaaaaaga ataaagaaga ggaaaggtga agctatggct ctaaatgaga 960
aaagaattct ggagaaagtg caaagtagat tcgtagttag tttagcctac gcttatgaaa 1020
ccaaagatgc cttgtgcttg gtgctcacca ttatgaatgg aggggatttg aagtttcaca 1080
tttacaacct gggcaatccc ggctttgatg agcagagagc cgttttctat gctgcagagc 1140
tgtgttgcgg cttggaagat ttacagaggg aaagaattgt atacagagac ttgaagcctg 1200
agaatattct ccttgatgat cgtggacaca tccggatttc agacctcggt ttggccacag 1260
agatcccaga aggacagagg gttcgaggaa gagttggaac agtcggctac atggcacctg 1320
aagttgtcaa taatgaaaag tatacgttta gtcccgattg gtggggactt ggctgtctga 1380
tctatgaaat gattcaggga cattctccat tcaaaaaata caaagagaaa gtcaaatggg 1440
aggaggtcga tcaaagaatc aagaatgata ccgaggagta ttctgagaag ttttcagagg 1500
atgccaaatc tatctgcagg atgttactca ccaagaatcc aagcaagcgg ctgggctgca 1560
ggggcgaggg agcggctggg gtgaagcagc accccgtgtt caaggacatc aacttcagga 1620
ggctggaggc aaacatgctg gagccccctt tctgtcctga tcctcatgcc gtttactgta 1680
aggacgtcct ggatatcgag cagttctcgg cggtgaaagg gatctacctg gacaccgcag 1740
atgaagactt ctatgctcgg tttgctaccg ggtgtgtctc catcccctgg cagaatgaga 1800
tgatcgaatc cgggtgtttc aaagacatca acaaaagtga aagtgaggaa gctttgccat 1860
tagatctaga caagaacata catacccccgg tttccagacc aaacagaggc ttcttctata 1920
gactcttcag aagaggggcc tgcctgacca tggtccccag tgagaaggaa gtggaaccca 1980
agcaatgctg agcaccccgg tgcggaccac agagcagacc ctggcgccag gaaggagcat 2040
gtgttagcgt ctcgtcccac ctggaattgt aataaataca tctaaataaa acatgccttg 2100
ggagtgtaca gac 2113

<210> 2
<211> 1737
<212> DNA
<213> Homo Sapiens

<220>
<221> CDS
<222> (1)...(1737)

<400> 2

```
atg gag ctc gag aac atc gtg gcc aac tcg ctg ctg ctg aaa gcg cgt   48
Met Glu Leu Glu Asn Ile Val Ala Asn Ser Leu Leu Leu Lys Ala Arg
 1               5                  10                  15

caa gga gga tat ggc aaa aaa agt ggt cgt agt aaa aaa tgg aag gag   96
Gln Gly Gly Tyr Gly Lys Lys Ser Gly Arg Ser Lys Lys Trp Lys Glu
            20                  25                  30

ata ctg aca ctg cct cct gtc agc cag tgc agt gag ctt aga cat tcc   144
Ile Leu Thr Leu Pro Pro Val Ser Gln Cys Ser Glu Leu Arg His Ser
        35                  40                  45

att gaa aag gat tat agc agt ctt tgt gac aag caa ccg ata gga aga   192
Ile Glu Lys Asp Tyr Ser Ser Leu Cys Asp Lys Gln Pro Ile Gly Arg
    50                  55                  60

cgt ctc ttc agg cag ttc tgt gat acc aaa ccc act cta aag agg cac   240
Arg Leu Phe Arg Gln Phe Cys Asp Thr Lys Pro Thr Leu Lys Arg His
65                  70                  75                  80

att gaa ttc ttg gat gca gtg gca gaa tat gaa gtt gcc gat gat gag   288
Ile Glu Phe Leu Asp Ala Val Ala Glu Tyr Glu Val Ala Asp Asp Glu
                85                  90                  95

gac cga agt gat tgt gga ctg tca atc tta gat aga ttc ttc aat gat   336
Asp Arg Ser Asp Cys Gly Leu Ser Ile Leu Asp Arg Phe Phe Asn Asp
            100                 105                 110

aag ttg gca gcc cct tta cca gaa ata cct cca gat gtt gtg aca gaa   384
Lys Leu Ala Ala Pro Leu Pro Glu Ile Pro Pro Asp Val Val Thr Glu
        115                 120                 125

tgt aga ttg gga ctg aag gag gag aac cct tcc aaa aaa gcc ttt gag   432
Cys Arg Leu Gly Leu Lys Glu Glu Asn Pro Ser Lys Lys Ala Phe Glu
    130                 135                 140

gaa tgt act aga gtt gcc cat aac tac cta aga ggg gaa cca ttt gaa   480
Glu Cys Thr Arg Val Ala His Asn Tyr Leu Arg Gly Glu Pro Phe Glu
145                 150                 155                 160

gaa tac caa gaa agc tca tat ttt tct cag ttt tta caa tgg aaa tgg   528
Glu Tyr Gln Glu Ser Ser Tyr Phe Ser Gln Phe Leu Gln Trp Lys Trp
                165                 170                 175

ctg gaa agg caa ccc gta aca aag aac aca ttt aga cat tac aga gtt   576
Leu Glu Arg Gln Pro Val Thr Lys Asn Thr Phe Arg His Tyr Arg Val
            180                 185                 190

cta gga aaa ggc gga ttt gga gag gtt tgc gcc tgt caa gtg cga gcc   624
Leu Gly Lys Gly Gly Phe Gly Glu Val Cys Ala Cys Gln Val Arg Ala
        195                 200                 205
```

```
aca gga aaa atg tat gcc tgc aaa aag cta caa aaa aaa aga ata aag   672
Thr Gly Lys Met Tyr Ala Cys Lys Lys Leu Gln Lys Lys Arg Ile Lys
    210                 215                 220

aag agg aaa ggt gaa gct atg gct cta aat gag aaa aga att ctg gag   720
Lys Arg Lys Gly Glu Ala Met Ala Leu Asn Glu Lys Arg Ile Leu Glu
225                 230                 235                 240

aaa gtg caa agt aga ttc gta gtt agt tta gcc tac gct tat gaa acc   768
Lys Val Gln Ser Arg Phe Val Val Ser Leu Ala Tyr Ala Tyr Glu Thr
                245                 250                 255

aaa gat gcc ttg tgc ttg gtg ctc acc att atg aat gga ggg gat ttg   816
Lys Asp Ala Leu Cys Leu Val Leu Thr Ile Met Asn Gly Gly Asp Leu
            260                 265                 270

aag ttt cac att tac aac ctg ggc aat ccc ggc ttt gat gag cag aga   864
Lys Phe His Ile Tyr Asn Leu Gly Asn Pro Gly Phe Asp Glu Gln Arg
        275                 280                 285

gcc gtt ttc tat gct gca gag ctg tgt tgc ggc ttg gaa gat tta cag   912
Ala Val Phe Tyr Ala Ala Glu Leu Cys Cys Gly Leu Glu Asp Leu Gln
    290                 295                 300

agg gaa aga att gta tac aga gac ttg aag cct gag aat att ctc ctt   960
Arg Glu Arg Ile Val Tyr Arg Asp Leu Lys Pro Glu Asn Ile Leu Leu
305                 310                 315                 320

gat gat cgt gga cac atc cgg att tca gac ctc ggt ttg gcc aca gag   1008
Asp Asp Arg Gly His Ile Arg Ile Ser Asp Leu Gly Leu Ala Thr Glu
                325                 330                 335

atc cca gaa gga cag agg gtt cga gga aga gtt gga aca gtc ggc tac   1056
Ile Pro Glu Gly Gln Arg Val Arg Gly Arg Val Gly Thr Val Gly Tyr
            340                 345                 350

atg gca cct gaa gtt gtc aat aat gaa aag tat acg ttt agt ccc gat   1104
Met Ala Pro Glu Val Val Asn Asn Glu Lys Tyr Thr Phe Ser Pro Asp
        355                 360                 365

tgg tgg gga ctt ggc tgt ctg atc tat gaa atg att cag gga cat tct   1152
Trp Trp Gly Leu Gly Cys Leu Ile Tyr Glu Met Ile Gln Gly His Ser
    370                 375                 380

cca ttc aaa aaa tac aaa gag aaa gtc aaa tgg gag gag gtc gat caa   1200
Pro Phe Lys Lys Tyr Lys Glu Lys Val Lys Trp Glu Glu Val Asp Gln
385                 390                 395                 400

aga atc aag aat gat acc gag gag tat tct gag aag ttt tca gag gat   1248
Arg Ile Lys Asn Asp Thr Glu Glu Tyr Ser Glu Lys Phe Ser Glu Asp
                405                 410                 415

gcc aaa tct atc tgc agg atg tta ctc acc aag aat cca agc aag cgg   1296
Ala Lys Ser Ile Cys Arg Met Leu Leu Thr Lys Asn Pro Ser Lys Arg
            420                 425                 430

ctg ggc tgc agg ggc gag gga gcg gct ggg gtg aag cag cac ccc gtg   1344
Leu Gly Cys Arg Gly Glu Gly Ala Ala Gly Val Lys Gln His Pro Val
        435                 440                 445

ttc aag gac atc aac ttc agg agg ctg gag gca aac atg ctg gag ccc   1392
```

```
Phe Lys Asp Ile Asn Phe Arg Arg Leu Glu Ala Asn Met Leu Glu Pro
    450                 455                 460

cct ttc tgt cct gat cct cat gcc gtt tac tgt aag gac gtc ctg gat   1440
Pro Phe Cys Pro Asp Pro His Ala Val Tyr Cys Lys Asp Val Leu Asp
465                 470                 475                 480

atc gag cag ttc tcg gcg gtg aaa ggg atc tac ctg gac acc gca gat   1488
Ile Glu Gln Phe Ser Ala Val Lys Gly Ile Tyr Leu Asp Thr Ala Asp
                485                 490                 495

gaa gac ttc tat gct cgg ttt gct acc ggg tgt gtc tcc atc ccc tgg   1536
Glu Asp Phe Tyr Ala Arg Phe Ala Thr Gly Cys Val Ser Ile Pro Trp
            500                 505                 510

cag aat gag atg atc gaa tcc ggg tgt ttc aaa gac atc aac aaa agt   1584
Gln Asn Glu Met Ile Glu Ser Gly Cys Phe Lys Asp Ile Asn Lys Ser
        515                 520                 525

gaa agt gag gaa gct ttg cca tta gat cta gac aag aac ata cat acc   1632
Glu Ser Glu Glu Ala Leu Pro Leu Asp Leu Asp Lys Asn Ile His Thr
    530                 535                 540

ccg gtt tcc aga cca aac aga ggc ttc ttc tat aga ctc ttc aga aga   1680
Pro Val Ser Arg Pro Asn Arg Gly Phe Phe Tyr Arg Leu Phe Arg Arg
545                 550                 555                 560

ggg ggc tgc ctg acc atg gtc ccc agt gag aag gaa gtg gaa ccc aag   1728
Gly Gly Cys Leu Thr Met Val Pro Ser Glu Lys Glu Val Glu Pro Lys
                565                 570                 575

caa tgc tga                                                      1737
Gln Cys  *
```

<210> 3
<211> 578
<212> PRT
<213> Homo Sapiens

<400> 3

```
Met Glu Leu Glu Asn Ile Val Ala Asn Ser Leu Leu Leu Lys Ala Arg
 1               5                  10                  15
Gln Gly Gly Tyr Gly Lys Lys Ser Gly Arg Ser Lys Lys Trp Lys Glu
            20                  25                  30
Ile Leu Thr Leu Pro Pro Val Ser Gln Cys Ser Glu Leu Arg His Ser
        35                  40                  45
Ile Glu Lys Asp Tyr Ser Ser Leu Cys Asp Lys Gln Pro Ile Gly Arg
    50                  55                  60
Arg Leu Phe Arg Gln Phe Cys Asp Thr Lys Pro Thr Leu Lys Arg His
65                  70                  75                  80
Ile Glu Phe Leu Asp Ala Val Ala Glu Tyr Glu Val Ala Asp Asp Glu
                85                  90                  95
Asp Arg Ser Asp Cys Gly Leu Ser Ile Leu Asp Arg Phe Phe Asn Asp
            100                 105                 110
Lys Leu Ala Ala Pro Leu Pro Glu Ile Pro Pro Asp Val Val Thr Glu
        115                 120                 125
Cys Arg Leu Gly Leu Lys Glu Glu Asn Pro Ser Lys Lys Ala Phe Glu
    130                 135                 140
Glu Cys Thr Arg Val Ala His Asn Tyr Leu Arg Gly Glu Pro Phe Glu
145                 150                 155                 160
```

```
Glu Tyr Gln Glu Ser Ser Tyr Phe Ser Gln Phe Leu Gln Trp Lys Trp
                165                 170                 175
Leu Glu Arg Gln Pro Val Thr Lys Asn Thr Phe Arg His Tyr Arg Val
            180                 185                 190
Leu Gly Lys Gly Gly Phe Gly Glu Val Cys Ala Cys Gln Val Arg Ala
        195                 200                 205
Thr Gly Lys Met Tyr Ala Cys Lys Lys Leu Gln Lys Lys Arg Ile Lys
    210                 215                 220
Lys Arg Lys Gly Glu Ala Met Ala Leu Asn Glu Lys Arg Ile Leu Glu
225                 230                 235                 240
Lys Val Gln Ser Arg Phe Val Val Ser Leu Ala Tyr Ala Tyr Glu Thr
                245                 250                 255
Lys Asp Ala Leu Cys Leu Val Leu Thr Ile Met Asn Gly Gly Asp Leu
            260                 265                 270
Lys Phe His Ile Tyr Asn Leu Gly Asn Pro Gly Phe Asp Glu Gln Arg
        275                 280                 285
Ala Val Phe Tyr Ala Ala Glu Leu Cys Cys Gly Leu Glu Asp Leu Gln
    290                 295                 300
Arg Glu Arg Ile Val Tyr Arg Asp Leu Lys Pro Glu Asn Ile Leu Leu
305                 310                 315                 320
Asp Asp Arg Gly His Ile Arg Ile Ser Asp Leu Gly Leu Ala Thr Glu
                325                 330                 335
Ile Pro Glu Gly Gln Arg Val Arg Gly Arg Val Gly Thr Val Gly Tyr
            340                 345                 350
Met Ala Pro Glu Val Val Asn Asn Glu Lys Tyr Thr Phe Ser Pro Asp
        355                 360                 365
Trp Trp Gly Leu Gly Cys Leu Ile Tyr Glu Met Ile Gln Gly His Ser
    370                 375                 380
Pro Phe Lys Lys Tyr Lys Glu Lys Val Lys Trp Glu Glu Val Asp Gln
385                 390                 395                 400
Arg Ile Lys Asn Asp Thr Glu Glu Tyr Ser Glu Lys Phe Ser Glu Asp
                405                 410                 415
Ala Lys Ser Ile Cys Arg Met Leu Leu Thr Lys Asn Pro Ser Lys Arg
            420                 425                 430
Leu Gly Cys Arg Gly Glu Gly Ala Ala Gly Val Lys Gln His Pro Val
        435                 440                 445
Phe Lys Asp Ile Asn Phe Arg Arg Leu Glu Ala Asn Met Leu Glu Pro
    450                 455                 460
Pro Phe Cys Pro Asp Pro His Ala Val Tyr Cys Lys Asp Val Leu Asp
465                 470                 475                 480
Ile Glu Gln Phe Ser Ala Val Lys Gly Ile Tyr Leu Asp Thr Ala Asp
                485                 490                 495
Glu Asp Phe Tyr Ala Arg Phe Ala Thr Gly Cys Val Ser Ile Pro Trp
            500                 505                 510
Gln Asn Glu Met Ile Glu Ser Gly Cys Phe Lys Asp Ile Asn Lys Ser
        515                 520                 525
Glu Ser Glu Glu Ala Leu Pro Leu Asp Leu Asp Lys Asn Ile His Thr
    530                 535                 540
Pro Val Ser Arg Pro Asn Arg Gly Phe Phe Tyr Arg Leu Phe Arg Arg
545                 550                 555                 560
Gly Gly Cys Leu Thr Met Val Pro Ser Glu Lys Glu Val Glu Pro Lys
                565                 570                 575
Gln Cys
```

<210> 4
<211> 1242
<212> DNA
<213> Homo Sapiens

<400> 4

```
aagaggctct ctctgcccat ggatatccgc ctgccccagg aattcctaca gaagctacag 60
atggagagcc cagatctgcc caagccgctc agccgcatgt cccgccgggc ctccctgtca 120
gacattggct ttgggaaact ggaaacatac gtgaaactgg acaaactggg agagggcacc 180
tatgccacag tcttcaaagg gcgcagcaaa ctgacggaga accttgtggc cctgaaagag 240
atccggctgg agcacgagga gggagcgccc tgcactgcca tccgagaggt gtctctgctg 300
aagaacctga agcacgccaa tattgtgacc ctgcatgacc tcatccacac agatcggtcc 360
ctcaccctgg tgtttgagta cctggacagt gacctgaagc agtatctgga ccactgtggg 420
aacctcatga gcatgcacaa cgtcaagatt ttcatgttcc agctgctccg gggcctcgcc 480
tactgtcaca cccgcaagat cctgcaccgg gacctgaagc cccagaacct gctcatcaac 540
gagaggggg agctgaagct ggccgacttt ggactggcca gggccaagtc agtgcccaca 600
aagacttact ccaatgaggt ggtgaccctg tggtacaggc cccccgatgt gctgctggga 660
tccacagagt actccacccc cattgctatg tggggcgtgg gctgcatcca ctacgagatg 720
gccacaggga ggcccctctt cccgggctcc acagtcaagg aggagctgca cctcatcttt 780
cgcctcctcg ggacccccac agaagagacg tggcccggcg tgaccgcctt ctctgagttc 840
cgcacctaca gcttcccctg ctacctcccg cagccgctca tcaaccacgc gcccaggttg 900
gatacggatg gcatccacct cctgagcagc ctgctcgtgt atgaatccaa gagtcgcatg 960
tcagcagagg ctgccctgag tcactcctac ttccggtctc tgggagagcg tgtgcaccag 1020
cttgaagaca ctgcctccat cttctccctg aaggagatcc agctccagaa ggacccaggc 1080
taccgaggct tggccttcca gcagccagga cgagggaaga acaggcggca gagcatcttc 1140
tgagccacgc ccaccttgct gtggccaagg gacaagagac cacatggagc acaaattcgg 1200
gtaggatgga gcctgtgtgg ccctcggagg actgaacaac ga                    1242

<210> 5
<211> 1125
<212> DNA
<213> Homo Sapiens

<220>
<221> CDS
<222> (1)...(1125)

<400> 5
atg gat atc cgc ctg ccc cag gaa ttc cta cag aag cta cag atg gag   48
Met Asp Ile Arg Leu Pro Gln Glu Phe Leu Gln Lys Leu Gln Met Glu
 1               5                  10                  15

agc cca gat ctg ccc aag ccg ctc agc cgc atg tcc cgc cgg gcc tcc   96
Ser Pro Asp Leu Pro Lys Pro Leu Ser Arg Met Ser Arg Arg Ala Ser
             20                  25                  30

ctg tca gac att ggc ttt ggg aaa ctg gaa aca tac gtg aaa ctg gac   144
Leu Ser Asp Ile Gly Phe Gly Lys Leu Glu Thr Tyr Val Lys Leu Asp
         35                  40                  45

aaa ctg gga gag ggc acc tat gcc aca gtc ttc aaa ggg cgc agc aaa   192
Lys Leu Gly Glu Gly Thr Tyr Ala Thr Val Phe Lys Gly Arg Ser Lys
     50                  55                  60

ctg acg gag aac ctt gtg gcc ctg aaa gag atc cgg ctg gag cac gag   240
Leu Thr Glu Asn Leu Val Ala Leu Lys Glu Ile Arg Leu Glu His Glu
 65                  70                  75                  80

gag gga gcg ccc tgc act gcc atc cga gag gtg tct ctg ctg aag aac   288
Glu Gly Ala Pro Cys Thr Ala Ile Arg Glu Val Ser Leu Leu Lys Asn
                 85                  90                  95

ctg aag cac gcc aat att gtg acc ctg cat gac ctc atc cac aca gat   336
Leu Lys His Ala Asn Ile Val Thr Leu His Asp Leu Ile His Thr Asp
            100                 105                 110

cgg tcc ctc acc ctg gtg ttt gag tac ctg gac agt gac ctg aag cag   384
Arg Ser Leu Thr Leu Val Phe Glu Tyr Leu Asp Ser Asp Leu Lys Gln
```

```
        115                 120                 125

tat ctg gac cac tgt ggg aac ctc atg agc atg cac aac gtc aag att   432
Tyr Leu Asp His Cys Gly Asn Leu Met Ser Met His Asn Val Lys Ile
    130                 135                 140

ttc atg ttc cag ctg ctc cgg ggc ctc gcc tac tgt cac acc cgc aag   480
Phe Met Phe Gln Leu Leu Arg Gly Leu Ala Tyr Cys His Thr Arg Lys
145                 150                 155                 160

atc ctg cac cgg gac ctg aag ccc cag aac ctg ctc atc aac gag agg   528
Ile Leu His Arg Asp Leu Lys Pro Gln Asn Leu Leu Ile Asn Glu Arg
                165                 170                 175

ggg gag ctg aag ctg gcc gac ttt gga ctg gcc agg gcc aag tca gtg   576
Gly Glu Leu Lys Leu Ala Asp Phe Gly Leu Ala Arg Ala Lys Ser Val
            180                 185                 190

ccc aca aag act tac tcc aat gag gtg gtg acc ctg tgg tac agg ccc   624
Pro Thr Lys Thr Tyr Ser Asn Glu Val Val Thr Leu Trp Tyr Arg Pro
        195                 200                 205

ccc gat gtg ctg ctg gga tcc aca gag tac tcc acc ccc att gct atg   672
Pro Asp Val Leu Leu Gly Ser Thr Glu Tyr Ser Thr Pro Ile Ala Met
    210                 215                 220

tgg ggc gtg ggc tgc atc cac tac gag atg gcc aca ggg agg ccc ctc   720
Trp Gly Val Gly Cys Ile His Tyr Glu Met Ala Thr Gly Arg Pro Leu
225                 230                 235                 240

ttc ccg ggc tcc aca gtc aag gag gag ctg cac ctc atc ttt cgc ctc   768
Phe Pro Gly Ser Thr Val Lys Glu Glu Leu His Leu Ile Phe Arg Leu
                245                 250                 255

ctc ggg acc ccc aca gaa gag acg tgg ccc ggc gtg acc gcc ttc tct   816
Leu Gly Thr Pro Thr Glu Glu Thr Trp Pro Gly Val Thr Ala Phe Ser
            260                 265                 270

gag ttc cgc acc tac agc ttc ccc tgc tac ctc ccg cag ccg ctc atc   864
Glu Phe Arg Thr Tyr Ser Phe Pro Cys Tyr Leu Pro Gln Pro Leu Ile
        275                 280                 285

aac cac gcg ccc agg ttg gat acg gat ggc atc cac ctc ctg agc agc   912
Asn His Ala Pro Arg Leu Asp Thr Asp Gly Ile His Leu Leu Ser Ser
    290                 295                 300

ctg ctc gtg tat gaa tcc aag agt cgc atg tca gca gag gct gcc ctg   960
Leu Leu Val Tyr Glu Ser Lys Ser Arg Met Ser Ala Glu Ala Ala Leu
305                 310                 315                 320

agt cac tcc tac ttc cgg tct ctg gga gag cgt gtg cac cag ctt gaa   1008
Ser His Ser Tyr Phe Arg Ser Leu Gly Glu Arg Val His Gln Leu Glu
                325                 330                 335

gac act gcc tcc atc ttc tcc ctg aag gag atc cag ctc cag aag gac   1056
Asp Thr Ala Ser Ile Phe Ser Leu Lys Glu Ile Gln Leu Gln Lys Asp
            340                 345                 350

cca ggc tac cga ggc ttg gcc ttc cag cag cca gga cga ggg aag aac   1104
Pro Gly Tyr Arg Gly Leu Ala Phe Gln Gln Pro Gly Arg Gly Lys Asn
        355                 360                 365
```

```
agg cgg cag agc atc ttc tga                                          1125
Arg Arg Gln Ser Ile Phe  *
    370


<210> 6
<211> 374
<212> PRT
<213> Homo Sapiens

<400> 6
Met Asp Ile Arg Leu Pro Gln Glu Phe Leu Gln Lys Leu Gln Met Glu
 1               5                  10                  15
Ser Pro Asp Leu Pro Lys Pro Leu Ser Arg Met Ser Arg Arg Ala Ser
            20                  25                  30
Leu Ser Asp Ile Gly Phe Gly Lys Leu Glu Thr Tyr Val Lys Leu Asp
        35                  40                  45
Lys Leu Gly Glu Gly Thr Tyr Ala Thr Val Phe Lys Gly Arg Ser Lys
    50                  55                  60
Leu Thr Glu Asn Leu Val Ala Leu Lys Glu Ile Arg Leu Glu His Glu
65                  70                  75                  80
Glu Gly Ala Pro Cys Thr Ala Ile Arg Glu Val Ser Leu Leu Lys Asn
                85                  90                  95
Leu Lys His Ala Asn Ile Val Thr Leu His Asp Leu Ile His Thr Asp
            100                 105                 110
Arg Ser Leu Thr Leu Val Phe Glu Tyr Leu Asp Ser Asp Leu Lys Gln
        115                 120                 125
Tyr Leu Asp His Cys Gly Asn Leu Met Ser Met His Asn Val Lys Ile
    130                 135                 140
Phe Met Phe Gln Leu Leu Arg Gly Leu Ala Tyr Cys His Thr Arg Lys
145                 150                 155                 160
Ile Leu His Arg Asp Leu Lys Pro Gln Asn Leu Leu Ile Asn Glu Arg
                165                 170                 175
Gly Glu Leu Lys Leu Ala Asp Phe Gly Leu Ala Arg Ala Lys Ser Val
            180                 185                 190
Pro Thr Lys Thr Tyr Ser Asn Glu Val Val Thr Leu Trp Tyr Arg Pro
        195                 200                 205
Pro Asp Val Leu Leu Gly Ser Thr Glu Tyr Ser Thr Pro Ile Ala Met
    210                 215                 220
Trp Gly Val Gly Cys Ile His Tyr Glu Met Ala Thr Gly Arg Pro Leu
225                 230                 235                 240
Phe Pro Gly Ser Thr Val Lys Glu Glu Leu His Leu Ile Phe Arg Leu
                245                 250                 255
Leu Gly Thr Pro Thr Glu Glu Thr Trp Pro Gly Val Thr Ala Phe Ser
            260                 265                 270
Glu Phe Arg Thr Tyr Ser Phe Pro Cys Tyr Leu Pro Gln Pro Leu Ile
        275                 280                 285
Asn His Ala Pro Arg Leu Asp Thr Asp Gly Ile His Leu Leu Ser Ser
    290                 295                 300
Leu Leu Val Tyr Glu Ser Lys Ser Arg Met Ser Ala Glu Ala Ala Leu
305                 310                 315                 320
Ser His Ser Tyr Phe Arg Ser Leu Gly Glu Arg Val His Gln Leu Glu
                325                 330                 335
Asp Thr Ala Ser Ile Phe Ser Leu Lys Glu Ile Gln Leu Gln Lys Asp
            340                 345                 350
Pro Gly Tyr Arg Gly Leu Ala Phe Gln Gln Pro Gly Arg Gly Lys Asn
        355                 360                 365
Arg Arg Gln Ser Ile Phe
    370


<210> 7
```

<211> 4151
<212> DNA
<213> Homo Sapiens

<400> 7
cgcaacttca tagacggttg tgaagacaag ctagacggca gatgaaagtt cttggcacag 60
agtgaacact tgataaacta tatggcagcc acagcctccg gagccgttgc acacctacct 120
gcccggccga cttacctgta cttgccgccg tcccggctca cctggcggtg cccgaggagt 180
agtcgctgga gtccgcgcct ccctgggact gcaatgtgcc gatcttagct gctgcctgag 240
aggatgtctg gggtgtccga gcccctgagt cgagtaaagt tgggcacgtt acgccggcct 300
gaaggccctg gagagcccat ggtggtggta ccagtagatg tggaaaagga ggacgtgcgt 360
atcctcaagg tctgcttcta tagcaacagc ttcaatcctg ggaaaaactt caaactggtc 420
aaatgcactg tccagacgga gatccgggag atcatcacct ccatcctgct gagcgggcgg 480
atcgggccca acatccggtt ggctgagtgc tatgggctga ggctgaagca catgaagtcc 540
gatgagatcc actggctgca cccacagatg acagtgggtg aggtgcagga caagtatgag 600
tgtctgcacg tggaagccga gtggaggtat gaccttcaaa tccgctactt gccagaagac 660
ttcatggaga gcctgaagga ggacaggacc acgctgctct atttttacca acagctccgg 720
aacgactaca tgcagcgcta cgccagcaag gtcagcgagg gcatggccct gcagctgggc 780
tgcctggagc tcaggcggtt cttcaaggat atgccccaca atgcacttga caagaagtcc 840
aacttcgagc tcctagaaaa ggaagtgggg ctggacttgt ttttcccaaa gcagatgcag 900
gagaacttaa agcccaaaca gttccggaag atgatccagc agaccttcca gcagtacgcc 960
tcgctcaggg aggaggagtg cgtcatgaag ttcttcaaca ctctcgccgg cttcgccaac 1020
atcgaccagg agacctaccg ctgtgaactc attcaaggat ggaacattac tgtggacctg 1080
gtcattggcc ctaaagggat ccgccagctg actagtcagg acgcaaagcc cacctgcctg 1140
gccgagttca agcagatcag gtccatcagg tgcctcccgc tggaggaggg ccaggcagta 1200
cttcagctgg gcattgaagg tgccccccag gccttgtcca tcaaaacctc atccctagca 1260
gaggctgaga acatggctga cctcatagac ggctactgcc ggctgcaggg tgagcaccaa 1320
ggctctctca tcatccatcc taggaaagat ggtgagaagc ggaacagcct gccccagatc 1380
cccatgctaa acctggaggc ccggcggtcc cacctctcag agagctgcag catagagtca 1440
gacatctacg cagagattcc cgacgaaacc ctgcgaaggc ccggaggtcc acagtatggc 1500
attgcccgtg aagatgtggt cctgaatcgt attcttgggg aaggcctttt tggggaggtc 1560
tatgaaggtg tctacacaaa tcataaaggg gagaaaatca atgtagctgt caagacctgc 1620
aagaaagact gcactctgga caacaaggag aagttcatga gcgaggcagt gatcatgaag 1680
aacctcgacc acccgcacat cgtgaagctg atcggcatca ttgaagagga gcccacctgg 1740
atcatcatgg aattgtatcc ctatggggag ctgggccact acctggagcg gaacaagaac 1800
tccctgaagg tgctcaccct cgtgctgtac tcactgcaga tatgcaaagc catggcctac 1860
ctggagagca tcaactgcgt gcacagggac attgctgtcc ggaacatcct ggtggcctcc 1920
cctgagtgtg tgaagctggg ggactttggt ctttcccggt acattgagga cgaggactat 1980
tacaaagcct ctgtgactcg tctccccatc aaatggatgt ccccagagtc cattaacttc 2040
cgacgcttca cgacagccag tgacgtctgg atgttcgccg tgtgcatgtg ggagatcctg 2100
agctttggga agcagccctt cttctggctg gagaacaagg atgtcatcgg ggtgctggag 2160
aaaggagacc ggctgcccaa gcctgatctc tgtccaccgg tcctttatac cctcatgacc 2220
cgctgctggg actacgaccc cagtgaccgg ccccgcttca ccgagctggt gtgcagcctc 2280
agtgacgttt atcagatgga gaaggacatt gccatggagc aagagaggaa tgctcgctac 2340
cgaaccccca aaatcttgga gcccacagcc ttccaggaac ccccacccaa gcccagccga 2400
cctaagtaca gacccccctcc gcaaaccaac ctcctggctc caaagctgca gttccaggtt 2460
cctgagggtc tgtgtgccag ctctcctacg ctcaccagcc ctatggagta tccatctccc 2520
gttaactcac tgcacacccc acctctccac cggcacaatg tcttcaaacg ccacagcatg 2580
cgggaggagg acttcatcca acccagcagc cgagaagagg cccagcagct gtgggaggct 2640
gaaaaggtca aaatgcggca aatcctggac aaacagcaga agcagatggt ggaggactac 2700
cagtggctca ggcaggagga gaagtccctg gaccccatgg tttatatgaa tgataagtcc 2760
ccattgacgc cagagaagga ggtcggctac ctggagttca cagggccccc acagaagccc 2820
ccgaggctgg gcgcacagtc catccagccc acagctaacc tggaccggac cgatgacctg 2880
gtgtacctca atgtcatgga gctggtgcgg gccgtgctgg agctcaagaa tgagctctgt 2940
cagctgcccc ccgagggcta cgtggtggtg gtgaagaatg tggggctgac cctgcggaag 3000
ctcatcggga gcgtggatga tctcctgcct tccttgccgt catcttcacg gacagagatc 3060
gagggcaccc agaaactgct caacaaagac ctggcagagc tcatcaacaa gatgcggctg 3120
gcgcagcaga acgccgtgac ctccctgagt gaggagtgca agaggcagat gctgacggct 3180
tcacacaccc tggctgtgga cgccaagaac ctgctcgacg ctgtggacca ggccaaggtt 3240
ctggccaatc tggcccaccc acctgcagag tgacggaggg tggggccacc tgcctgcgtc 3300
ttcgcccctg cctgccatgt acctcccctg ccttgctgtt ggtcatgtgg gtcttccagg 3360
gagaaggcca aggggagtca ccttcccttg ccactttgca cgacgccctc tccccacccc 3420

```
tacccctggc tgtactgctc aggctgcagc tggacagagg ggactctggg ctatggacac 3480
agggtgacgg tgacaaagat ggctcagagg gggactgctg ctgcctggcc actgctccct 3540
aagccagcct ggtccatgca gggggctcct gggggtgggg aggtgtcaca tggtgcccct 3600
agctttatat atggacatgg caggccgatt tgggaaccaa gctattcctt tcccttcctc 3660
ttcggccctc agatgtccct tgatgcacag agaagctggg gaggagcttt gttttggggg 3720
tcaggcagcc agtgagatga gggatgggcc tggcattctt gtacagtgta tattgaaatt 3780
tatttaatgt gagtttggtc tggactgaca gcatgtgccc tcctgaggga ggacctgggg 3840
cacagtccag gaacaagcta attgggagtc caggcacagg atgctgtgtt gtcaacaaac 3900
caagcatcag ggggaagaag cagagagatg cggccaagat aggaccttgg gccaaatccg 3960
ctctcttcct gcccctcttt ctctttcttc ctttactttc ccttgctttt ccctcttttc 4020
ttactcctcc tctttctctc cccaaccccc attctcatct gcacccttct tttctcatgt 4080
gtttgcataa acattctttt aacttctttc tatttgactt gtggttgaat taaaattgtc 4140
ccatttgctt g                                                      4151

<210> 8
<211> 3030
<212> DNA
<213> Homo Sapiens

<220>
<221> CDS
<222> (1)...(3030)

<400> 8
atg tct ggg gtg tcc gag ccc ctg agt cga gta aag ttg ggc acg tta   48
Met Ser Gly Val Ser Glu Pro Leu Ser Arg Val Lys Leu Gly Thr Leu
 1               5                  10                  15

cgc cgg cct gaa ggc cct gga gag ccc atg gtg gtg gta cca gta gat   96
Arg Arg Pro Glu Gly Pro Gly Glu Pro Met Val Val Val Pro Val Asp
            20                  25                  30

gtg gaa aag gag gac gtg cgt atc ctc aag gtc tgc ttc tat agc aac   144
Val Glu Lys Glu Asp Val Arg Ile Leu Lys Val Cys Phe Tyr Ser Asn
        35                  40                  45

agc ttc aat cct ggg aaa aac ttc aaa ctg gtc aaa tgc act gtc cag   192
Ser Phe Asn Pro Gly Lys Asn Phe Lys Leu Val Lys Cys Thr Val Gln
    50                  55                  60

acg gag atc cgg gag atc atc acc tcc atc ctg ctg agc ggg cgg atc   240
Thr Glu Ile Arg Glu Ile Ile Thr Ser Ile Leu Leu Ser Gly Arg Ile
65                  70                  75                  80

ggg ccc aac atc cgg ttg gct gag tgc tat ggg ctg agg ctg aag cac   288
Gly Pro Asn Ile Arg Leu Ala Glu Cys Tyr Gly Leu Arg Leu Lys His
                85                  90                  95

atg aag tcc gat gag atc cac tgg ctg cac cca cag atg aca gtg ggt   336
Met Lys Ser Asp Glu Ile His Trp Leu His Pro Gln Met Thr Val Gly
            100                 105                 110

gag gtg cag gac aag tat gag tgt ctg cac gtg gaa gcc gag tgg agg   384
Glu Val Gln Asp Lys Tyr Glu Cys Leu His Val Glu Ala Glu Trp Arg
        115                 120                 125

tat gac ctt caa atc cgc tac ttg cca gaa gac ttc atg gag agc ctg   432
Tyr Asp Leu Gln Ile Arg Tyr Leu Pro Glu Asp Phe Met Glu Ser Leu
    130                 135                 140

aag gag gac agg acc acg ctg ctc tat ttt tac caa cag ctc cgg aac   480
Lys Glu Asp Arg Thr Thr Leu Leu Tyr Phe Tyr Gln Gln Leu Arg Asn
```

```
145                 150                 155                 160

gac tac atg cag cgc tac gcc agc aag gtc agc gag ggc atg gcc ctg   528
Asp Tyr Met Gln Arg Tyr Ala Ser Lys Val Ser Glu Gly Met Ala Leu
                165                 170                 175

cag ctg ggc tgc ctg gag ctc agg cgg ttc ttc aag gat atg ccc cac   576
Gln Leu Gly Cys Leu Glu Leu Arg Arg Phe Phe Lys Asp Met Pro His
            180                 185                 190

aat gca ctt gac aag aag tcc aac ttc gag ctc cta gaa aag gaa gtg   624
Asn Ala Leu Asp Lys Lys Ser Asn Phe Glu Leu Leu Glu Lys Glu Val
        195                 200                 205

ggg ctg gac ttg ttt ttc cca aag cag atg cag gag aac tta aag ccc   672
Gly Leu Asp Leu Phe Phe Pro Lys Gln Met Gln Glu Asn Leu Lys Pro
    210                 215                 220

aaa cag ttc cgg aag atg atc cag cag acc ttc cag cag tac gcc tcg   720
Lys Gln Phe Arg Lys Met Ile Gln Gln Thr Phe Gln Gln Tyr Ala Ser
225                 230                 235                 240

ctc agg gag gag gag tgc gtc atg aag ttc ttc aac act ctc gcc ggc   768
Leu Arg Glu Glu Glu Cys Val Met Lys Phe Phe Asn Thr Leu Ala Gly
                245                 250                 255

ttc gcc aac atc gac cag gag acc tac cgc tgt gaa ctc att caa gga   816
Phe Ala Asn Ile Asp Gln Glu Thr Tyr Arg Cys Glu Leu Ile Gln Gly
            260                 265                 270

tgg aac att act gtg gac ctg gtc att ggc cct aaa ggg atc cgc cag   864
Trp Asn Ile Thr Val Asp Leu Val Ile Gly Pro Lys Gly Ile Arg Gln
        275                 280                 285

ctg act agt cag gac gca aag ccc acc tgc ctg gcc gag ttc aag cag   912
Leu Thr Ser Gln Asp Ala Lys Pro Thr Cys Leu Ala Glu Phe Lys Gln
    290                 295                 300

atc agg tcc atc agg tgc ctc ccg ctg gag gag ggc cag gca gta ctt   960
Ile Arg Ser Ile Arg Cys Leu Pro Leu Glu Glu Gly Gln Ala Val Leu
305                 310                 315                 320

cag ctg ggc att gaa ggt gcc ccc cag gcc ttg tcc atc aaa acc tca  1008
Gln Leu Gly Ile Glu Gly Ala Pro Gln Ala Leu Ser Ile Lys Thr Ser
                325                 330                 335

tcc cta gca gag gct gag aac atg gct gac ctc ata gac ggc tac tgc  1056
Ser Leu Ala Glu Ala Glu Asn Met Ala Asp Leu Ile Asp Gly Tyr Cys
            340                 345                 350

cgg ctg cag ggt gag cac caa ggc tct ctc atc atc cat cct agg aaa  1104
Arg Leu Gln Gly Glu His Gln Gly Ser Leu Ile Ile His Pro Arg Lys
        355                 360                 365

gat ggt gag aag cgg aac agc ctg ccc cag atc ccc atg cta aac ctg  1152
Asp Gly Glu Lys Arg Asn Ser Leu Pro Gln Ile Pro Met Leu Asn Leu
    370                 375                 380

gag gcc cgg cgg tcc cac ctc tca gag agc tgc agc ata gag tca gac  1200
Glu Ala Arg Arg Ser His Leu Ser Glu Ser Cys Ser Ile Glu Ser Asp
385                 390                 395                 400
```

```
atc tac gca gag att ccc gac gaa acc ctg cga agg ccc gga ggt cca   1248
Ile Tyr Ala Glu Ile Pro Asp Glu Thr Leu Arg Arg Pro Gly Gly Pro
                405                 410                 415

cag tat ggc att gcc cgt gaa gat gtg gtc ctg aat cgt att ctt ggg   1296
Gln Tyr Gly Ile Ala Arg Glu Asp Val Val Leu Asn Arg Ile Leu Gly
            420                 425                 430

gaa ggc ctt ttt ggg gag gtc tat gaa ggt gtc tac aca aat cat aaa   1344
Glu Gly Leu Phe Gly Glu Val Tyr Glu Gly Val Tyr Thr Asn His Lys
        435                 440                 445

ggg gag aaa atc aat gta gct gtc aag acc tgc aag aaa gac tgc act   1392
Gly Glu Lys Ile Asn Val Ala Val Lys Thr Cys Lys Lys Asp Cys Thr
    450                 455                 460

ctg gac aac aag gag aag ttc atg agc gag gca gtg atc atg aag aac   1440
Leu Asp Asn Lys Glu Lys Phe Met Ser Glu Ala Val Ile Met Lys Asn
465                 470                 475                 480

ctc gac cac ccg cac atc gtg aag ctg atc ggc atc att gaa gag gag   1488
Leu Asp His Pro His Ile Val Lys Leu Ile Gly Ile Ile Glu Glu Glu
                485                 490                 495

ccc acc tgg atc atc atg gaa ttg tat ccc tat ggg gag ctg ggc cac   1536
Pro Thr Trp Ile Ile Met Glu Leu Tyr Pro Tyr Gly Glu Leu Gly His
            500                 505                 510

tac ctg gag cgg aac aag aac tcc ctg aag gtg ctc acc ctc gtg ctg   1584
Tyr Leu Glu Arg Asn Lys Asn Ser Leu Lys Val Leu Thr Leu Val Leu
        515                 520                 525

tac tca ctg cag ata tgc aaa gcc atg gcc tac ctg gag agc atc aac   1632
Tyr Ser Leu Gln Ile Cys Lys Ala Met Ala Tyr Leu Glu Ser Ile Asn
    530                 535                 540

tgc gtg cac agg gac att gct gtc cgg aac atc ctg gtg gcc tcc cct   1680
Cys Val His Arg Asp Ile Ala Val Arg Asn Ile Leu Val Ala Ser Pro
545                 550                 555                 560

gag tgt gtg aag ctg ggg gac ttt ggt ctt tcc cgg tac att gag gac   1728
Glu Cys Val Lys Leu Gly Asp Phe Gly Leu Ser Arg Tyr Ile Glu Asp
                565                 570                 575

gag gac tat tac aaa gcc tct gtg act cgt ctc ccc atc aaa tgg atg   1776
Glu Asp Tyr Tyr Lys Ala Ser Val Thr Arg Leu Pro Ile Lys Trp Met
            580                 585                 590

tcc cca gag tcc att aac ttc cga cgc ttc acg aca gcc agt gac gtc   1824
Ser Pro Glu Ser Ile Asn Phe Arg Arg Phe Thr Thr Ala Ser Asp Val
        595                 600                 605

tgg atg ttc gcc gtg tgc atg tgg gag atc ctg agc ttt ggg aag cag   1872
Trp Met Phe Ala Val Cys Met Trp Glu Ile Leu Ser Phe Gly Lys Gln
    610                 615                 620

ccc ttc ttc tgg ctg gag aac aag gat gtc atc ggg gtg ctg gag aaa   1920
Pro Phe Phe Trp Leu Glu Asn Lys Asp Val Ile Gly Val Leu Glu Lys
625                 630                 635                 640

gga gac cgg ctg ccc aag cct gat ctc tgt cca ccg gtc ctt tat acc   1968
Gly Asp Arg Leu Pro Lys Pro Asp Leu Cys Pro Pro Val Leu Tyr Thr
```

```
                645                 650                 655

ctc atg acc cgc tgc tgg gac tac gac ccc agt gac cgg ccc cgc ttc   2016
Leu Met Thr Arg Cys Trp Asp Tyr Asp Pro Ser Asp Arg Pro Arg Phe
            660                 665                 670

acc gag ctg gtg tgc agc ctc agt gac gtt tat cag atg gag aag gac   2064
Thr Glu Leu Val Cys Ser Leu Ser Asp Val Tyr Gln Met Glu Lys Asp
        675                 680                 685

att gcc atg gag caa gag agg aat gct cgc tac cga acc ccc aaa atc   2112
Ile Ala Met Glu Gln Glu Arg Asn Ala Arg Tyr Arg Thr Pro Lys Ile
    690                 695                 700

ttg gag ccc aca gcc ttc cag gaa ccc cca ccc aag ccc agc cga cct   2160
Leu Glu Pro Thr Ala Phe Gln Glu Pro Pro Pro Lys Pro Ser Arg Pro
705                 710                 715                 720

aag tac aga ccc cct ccg caa acc aac ctc ctg gct cca aag ctg cag   2208
Lys Tyr Arg Pro Pro Pro Gln Thr Asn Leu Leu Ala Pro Lys Leu Gln
                725                 730                 735

ttc cag gtt cct gag ggt ctg tgt gcc agc tct cct acg ctc acc agc   2256
Phe Gln Val Pro Glu Gly Leu Cys Ala Ser Ser Pro Thr Leu Thr Ser
            740                 745                 750

cct atg gag tat cca tct ccc gtt aac tca ctg cac acc cca cct ctc   2304
Pro Met Glu Tyr Pro Ser Pro Val Asn Ser Leu His Thr Pro Pro Leu
        755                 760                 765

cac cgg cac aat gtc ttc aaa cgc cac agc atg cgg gag gag gac ttc   2352
His Arg His Asn Val Phe Lys Arg His Ser Met Arg Glu Glu Asp Phe
    770                 775                 780

atc caa ccc agc agc cga gaa gag gcc cag cag ctg tgg gag gct gaa   2400
Ile Gln Pro Ser Ser Arg Glu Glu Ala Gln Gln Leu Trp Glu Ala Glu
785                 790                 795                 800

aag gtc aaa atg cgg caa atc ctg gac aaa cag cag aag cag atg gtg   2448
Lys Val Lys Met Arg Gln Ile Leu Asp Lys Gln Gln Lys Gln Met Val
                805                 810                 815

gag gac tac cag tgg ctc agg cag gag gag aag tcc ctg gac ccc atg   2496
Glu Asp Tyr Gln Trp Leu Arg Gln Glu Glu Lys Ser Leu Asp Pro Met
            820                 825                 830

gtt tat atg aat gat aag tcc cca ttg acg cca gag aag gag gtc ggc   2544
Val Tyr Met Asn Asp Lys Ser Pro Leu Thr Pro Glu Lys Glu Val Gly
        835                 840                 845

tac ctg gag ttc aca ggg ccc cca cag aag ccc ccg agg ctg ggc gca   2592
Tyr Leu Glu Phe Thr Gly Pro Pro Gln Lys Pro Pro Arg Leu Gly Ala
    850                 855                 860

cag tcc atc cag ccc aca gct aac ctg gac cgg acc gat gac ctg gtg   2640
Gln Ser Ile Gln Pro Thr Ala Asn Leu Asp Arg Thr Asp Asp Leu Val
865                 870                 875                 880

tac ctc aat gtc atg gag ctg gtg cgg gcc gtg ctg gag ctc aag aat   2688
Tyr Leu Asn Val Met Glu Leu Val Arg Ala Val Leu Glu Leu Lys Asn
                885                 890                 895
```

```
gag ctc tgt cag ctg ccc ccc gag ggc tac gtg gtg gtg gtg aag aat   2736
Glu Leu Cys Gln Leu Pro Pro Glu Gly Tyr Val Val Val Val Lys Asn
            900                 905                 910

gtg ggg ctg acc ctg cgg aag ctc atc ggg agc gtg gat gat ctc ctg   2784
Val Gly Leu Thr Leu Arg Lys Leu Ile Gly Ser Val Asp Asp Leu Leu
        915                 920                 925

cct tcc ttg ccg tca tct tca cgg aca gag atc gag ggc acc cag aaa   2832
Pro Ser Leu Pro Ser Ser Ser Arg Thr Glu Ile Glu Gly Thr Gln Lys
    930                 935                 940

ctg ctc aac aaa gac ctg gca gag ctc atc aac aag atg cgg ctg gcg   2880
Leu Leu Asn Lys Asp Leu Ala Glu Leu Ile Asn Lys Met Arg Leu Ala
945                 950                 955                 960

cag cag aac gcc gtg acc tcc ctg agt gag gag tgc aag agg cag atg   2928
Gln Gln Asn Ala Val Thr Ser Leu Ser Glu Glu Cys Lys Arg Gln Met
                965                 970                 975

ctg acg gct tca cac acc ctg gct gtg gac gcc aag aac ctg ctc gac   2976
Leu Thr Ala Ser His Thr Leu Ala Val Asp Ala Lys Asn Leu Leu Asp
            980                 985                 990

gct gtg gac cag gcc aag gtt ctg gcc aat ctg gcc cac cca cct gca   3024
Ala Val Asp Gln Ala Lys Val Leu Ala Asn Leu Ala His Pro Pro Ala
         995                 1000                1005

gag tga                                                           3030
Glu  *
```

<210> 9
<211> 1009
<212> PRT
<213> Homo Sapiens

<400> 9

```
Met Ser Gly Val Ser Glu Pro Leu Ser Arg Val Lys Leu Gly Thr Leu
 1               5                  10                  15
Arg Arg Pro Glu Gly Pro Gly Glu Pro Met Val Val Val Pro Val Asp
            20                  25                  30
Val Glu Lys Glu Asp Val Arg Ile Leu Lys Val Cys Phe Tyr Ser Asn
        35                  40                  45
Ser Phe Asn Pro Gly Lys Asn Phe Lys Leu Val Lys Cys Thr Val Gln
    50                  55                  60
Thr Glu Ile Arg Glu Ile Ile Thr Ser Ile Leu Leu Ser Gly Arg Ile
65                  70                  75                  80
Gly Pro Asn Ile Arg Leu Ala Glu Cys Tyr Gly Leu Arg Leu Lys His
                85                  90                  95
Met Lys Ser Asp Glu Ile His Trp Leu His Pro Gln Met Thr Val Gly
            100                 105                 110
Glu Val Gln Asp Lys Tyr Glu Cys Leu His Val Glu Ala Glu Trp Arg
        115                 120                 125
Tyr Asp Leu Gln Ile Arg Tyr Leu Pro Glu Asp Phe Met Glu Ser Leu
    130                 135                 140
Lys Glu Asp Arg Thr Thr Leu Leu Tyr Phe Tyr Gln Gln Leu Arg Asn
145                 150                 155                 160
Asp Tyr Met Gln Arg Tyr Ala Ser Lys Val Ser Glu Gly Met Ala Leu
                165                 170                 175
Gln Leu Gly Cys Leu Glu Leu Arg Arg Phe Phe Lys Asp Met Pro His
```

```
            180                 185                 190
Asn Ala Leu Asp Lys Lys Ser Asn Phe Glu Leu Leu Glu Lys Glu Val
        195                 200                 205
Gly Leu Asp Leu Phe Phe Pro Lys Gln Met Gln Glu Asn Leu Lys Pro
    210                 215                 220
Lys Gln Phe Arg Lys Met Ile Gln Gln Thr Phe Gln Gln Tyr Ala Ser
225                 230                 235                 240
Leu Arg Glu Glu Glu Cys Val Met Lys Phe Phe Asn Thr Leu Ala Gly
                245                 250                 255
Phe Ala Asn Ile Asp Gln Glu Thr Tyr Arg Cys Glu Leu Ile Gln Gly
            260                 265                 270
Trp Asn Ile Thr Val Asp Leu Val Ile Gly Pro Lys Gly Ile Arg Gln
        275                 280                 285
Leu Thr Ser Gln Asp Ala Lys Pro Thr Cys Leu Ala Glu Phe Lys Gln
    290                 295                 300
Ile Arg Ser Ile Arg Cys Leu Pro Leu Glu Glu Gly Gln Ala Val Leu
305                 310                 315                 320
Gln Leu Gly Ile Glu Gly Ala Pro Gln Ala Leu Ser Ile Lys Thr Ser
                325                 330                 335
Ser Leu Ala Glu Ala Glu Asn Met Ala Asp Leu Ile Asp Gly Tyr Cys
            340                 345                 350
Arg Leu Gln Gly Glu His Gln Gly Ser Leu Ile Ile His Pro Arg Lys
        355                 360                 365
Asp Gly Glu Lys Arg Asn Ser Leu Pro Gln Ile Pro Met Leu Asn Leu
    370                 375                 380
Glu Ala Arg Arg Ser His Leu Ser Glu Ser Cys Ser Ile Glu Ser Asp
385                 390                 395                 400
Ile Tyr Ala Glu Ile Pro Asp Glu Thr Leu Arg Arg Pro Gly Gly Pro
                405                 410                 415
Gln Tyr Gly Ile Ala Arg Glu Asp Val Val Leu Asn Arg Ile Leu Gly
            420                 425                 430
Glu Gly Leu Phe Gly Glu Val Tyr Glu Gly Val Tyr Thr Asn His Lys
        435                 440                 445
Gly Glu Lys Ile Asn Val Ala Val Lys Thr Cys Lys Lys Asp Cys Thr
    450                 455                 460
Leu Asp Asn Lys Glu Lys Phe Met Ser Glu Ala Val Ile Met Lys Asn
465                 470                 475                 480
Leu Asp His Pro His Ile Val Lys Leu Ile Gly Ile Ile Glu Glu Glu
                485                 490                 495
Pro Thr Trp Ile Ile Met Glu Leu Tyr Pro Tyr Gly Glu Leu Gly His
            500                 505                 510
Tyr Leu Glu Arg Asn Lys Asn Ser Leu Lys Val Leu Thr Leu Val Leu
        515                 520                 525
Tyr Ser Leu Gln Ile Cys Lys Ala Met Ala Tyr Leu Glu Ser Ile Asn
    530                 535                 540
Cys Val His Arg Asp Ile Ala Val Arg Asn Ile Leu Val Ala Ser Pro
545                 550                 555                 560
Glu Cys Val Lys Leu Gly Asp Phe Gly Leu Ser Arg Tyr Ile Glu Asp
                565                 570                 575
Glu Asp Tyr Tyr Lys Ala Ser Val Thr Arg Leu Pro Ile Lys Trp Met
            580                 585                 590
Ser Pro Glu Ser Ile Asn Phe Arg Arg Phe Thr Thr Ala Ser Asp Val
        595                 600                 605
Trp Met Phe Ala Val Cys Met Trp Glu Ile Leu Ser Phe Gly Lys Gln
    610                 615                 620
Pro Phe Phe Trp Leu Glu Asn Lys Asp Val Ile Gly Val Leu Glu Lys
625                 630                 635                 640
Gly Asp Arg Leu Pro Lys Pro Asp Leu Cys Pro Pro Val Leu Tyr Thr
                645                 650                 655
Leu Met Thr Arg Cys Trp Asp Tyr Asp Pro Ser Asp Arg Pro Arg Phe
            660                 665                 670
Thr Glu Leu Val Cys Ser Leu Ser Asp Val Tyr Gln Met Glu Lys Asp
```

```
        675                 680                 685
Ile Ala Met Glu Gln Glu Arg Asn Ala Arg Tyr Arg Thr Pro Lys Ile
    690                 695                 700
Leu Glu Pro Thr Ala Phe Gln Glu Pro Pro Pro Lys Pro Ser Arg Pro
705                 710                 715                 720
Lys Tyr Arg Pro Pro Pro Gln Thr Asn Leu Leu Ala Pro Lys Leu Gln
                725                 730                 735
Phe Gln Val Pro Glu Gly Leu Cys Ala Ser Ser Pro Thr Leu Thr Ser
            740                 745                 750
Pro Met Glu Tyr Pro Ser Pro Val Asn Ser Leu His Thr Pro Pro Leu
        755                 760                 765
His Arg His Asn Val Phe Lys Arg His Ser Met Arg Glu Glu Asp Phe
    770                 775                 780
Ile Gln Pro Ser Ser Arg Glu Glu Ala Gln Gln Leu Trp Glu Ala Glu
785                 790                 795                 800
Lys Val Lys Met Arg Gln Ile Leu Asp Lys Gln Gln Lys Gln Met Val
                805                 810                 815
Glu Asp Tyr Gln Trp Leu Arg Gln Glu Glu Lys Ser Leu Asp Pro Met
            820                 825                 830
Val Tyr Met Asn Asp Lys Ser Pro Leu Thr Pro Glu Lys Glu Val Gly
        835                 840                 845
Tyr Leu Glu Phe Thr Gly Pro Pro Gln Lys Pro Pro Arg Leu Gly Ala
    850                 855                 860
Gln Ser Ile Gln Pro Thr Ala Asn Leu Asp Arg Thr Asp Asp Leu Val
865                 870                 875                 880
Tyr Leu Asn Val Met Glu Leu Val Arg Ala Val Leu Glu Leu Lys Asn
                885                 890                 895
Glu Leu Cys Gln Leu Pro Pro Glu Gly Tyr Val Val Val Val Lys Asn
            900                 905                 910
Val Gly Leu Thr Leu Arg Lys Leu Ile Gly Ser Val Asp Asp Leu Leu
        915                 920                 925
Pro Ser Leu Pro Ser Ser Ser Arg Thr Glu Ile Glu Gly Thr Gln Lys
    930                 935                 940
Leu Leu Asn Lys Asp Leu Ala Glu Leu Ile Asn Lys Met Arg Leu Ala
945                 950                 955                 960
Gln Gln Asn Ala Val Thr Ser Leu Ser Glu Glu Cys Lys Arg Gln Met
                965                 970                 975
Leu Thr Ala Ser His Thr Leu Ala Val Asp Ala Lys Asn Leu Leu Asp
            980                 985                 990
Ala Val Asp Gln Ala Lys Val Leu Ala Asn Leu Ala His Pro Pro Ala
        995                 1000                1005
Glu
```

<210> 10
<211> 6574
<212> DNA
<213> Homo Sapiens

<400> 10

```
agagccgcca gtggggagat gttgaagttc aaatatggag cgcggaatcc tttggatgct 60
ggtgctgctg aacccattgc cagccgggcc tccaggctga atctgttctt ccaggggaaa 120
ccacccttta tgactcaaca gcagatgtct cctctttccc gagaagggat attagatgcc 180
ctctttgttc tctttgaaga atgcagtcag cctgctctga tgaagattaa gcacgtgagc 240
aactttgtcc ggaagtattc cgacaccata gctgagttac aggagctcca gccttcggca 300
aaggacttcg aagtcagaag tcttgtaggt tgtggtcact ttgctgaagt gcaggtggta 360
agagagaaag caaccgggga catctatgct atgaaagtga tgaagaagaa ggctttattg 420
gcccaggagc aggtttcatt ttttgaggaa gagcggaaca tattatctcg aagcacaagc 480
ccgtggatcc cccaattaca gtatgccttt caggacaaaa atcaccttta tctgatggag 540
gaatatcagc ctggagggga cttgctgtca cttttgaata gatatgagga ccagttagat 600
gaaaacctga tacagtttta cctagctgag ctgattttgg ctgttcacag cgttcatctg 660
```

atgggatacg tgcatcgaga catcaagcct gagaacattc tcgttgaccg cacaggacac 720
atcaagctgg tggattttgg atctgccgcg aaaatgaatt caaacaagat ggtgaatgcc 780
aaactcccga ttgggacccc agattacatg gctcctgaag tgctgactgt gatgaacggg 840
gatggaaaag gcacctacgg cctggactgt gactggtggt cagtgggcgt gattgcctat 900
gagatgattt atgggagatc ccccttcgca gagggaacct ctgccagaac cttcaataac 960
attatgaatt tccagcggtt tttgaaattt ccagatgacc ccaaagtgag cagtgacttt 1020
cttgatctga ttcaaagctt gttgtgcggc cagaaagaga gactgaagtt tgaaggtctt 1080
tgctgccatc ctttcttctc taaaattgac tggaacaaca ttcgtaactc tcctcccccc 1140
ttcgttccca ccctcaagtc tgacgatgac acctccaatt ttgatgaacc agagaagaat 1200
tcgtgggttt catcctctcc gtgccagctg agcccctcag gcttctcggg tgaagaactg 1260
ccgtttgtgg ggttttcgta cagcaaggca ctggggattc ttggtagatc tgagtctgtt 1320
gtgtcgggtc tggactcccc tgccaagact agctccatgg aaaagaaact tctcatcaaa 1380
agcaaagagc tacaagactc tcaggacaag tgtcacaaga tggagcagga aatgacccgg 1440
ttacatcgga gagtgtcaga ggtggaggct gtgcttagtc agaaggaggt ggagctgaag 1500
gcctctgaga ctcagagatc cctcctggag caggaccttg ctacctacat cacagaatgc 1560
agtagcttaa agcgaagttt ggagcaagca cggatggagg tgtcccagga ggatgacaaa 1620
gcactgcagc ttctccatga tatcagagag cagagccgga agctccaaga aatcaaagag 1680
caggagtacc aggctcaagt ggaagaaatg aggttgatga tgaatcagtt ggaagaggat 1740
cttgtctcag caagaagacg gagtgatctc tacgaatctg agctgagaga gtctcggctt 1800
gctgctgaag aattcaagcg gaaagcgaca gaatgtcagc ataaactgtt gaaggctaag 1860
gatcaaggga agcctgaagt gggagaatat gcgaaactgg agaagatcaa tgctgagcag 1920
cagctcaaaa ttcaggagct ccaagagaaa ctggagaagg ctgcaaagga gcgagccgag 1980
agggagctgg agaagctgca gaaccgagag gattcttctg aaggcatcag aaagaagctg 2040
gtggaagctg aggaacgccg ccattctctg gagaacaagg taaagagact agagaccatg 2100
gagcgtagag aaaacagact gaaggatgac atccagacaa aatcccaaca gatccagcag 2160
atggctgata aaattctgga gctcgaagag aaacatcggg aggcccaagt ctcagcccag 2220
cacctagaag tgcacctgaa acagaaagag cagcactatg aggaaaagat taaagtgttg 2280
gacaatcaga taaagaaaga cctggctgac aaggagacac tggagaacat gatgcagaga 2340
cacgaggagg aggcccatga gaagggcaaa attctcagcg aacagaaggc gatgatcaat 2400
gctatggatt ccaagatcag atccctggaa cagaggattg tggaactgtc tgaagccaat 2460
aaacttgcag caaatagcag tctttttacc caaaggaaca tgaaggccca agaagagatg 2520
atttctgaac tcaggcaaca gaaattttac ctggagacac aggctgggaa gttggaggcc 2580
cagaaccgaa aactggagga gcagctggag aagatcagcc accaagacca cagtgacaag 2640
aatcggctgc tggaactgga gacaagattg cgggaggtca gtctagagca cgaggagcag 2700
aaactggagc tcaagcgcca gctcacagag ctacagctct ccctgcagga gcgcgagtca 2760
cagttgacag ccctgcaggc tgcacgggcg gccctggaga gccagcttcg ccaggcgaag 2820
acagagctgg aagagaccac agcagaagct gaagaggaga tccaggcact cacggcacat 2880
agagatgaaa tccagcgcaa atttgatgct cttcgtaaca gctgtactgt aatcacagac 2940
ctggaggagc agctaaacca gctgaccgag gacaacgctg aactcaacaa ccaaaacttc 3000
tacttgtcca aacaactcga tgaggcttct ggcgccaacg acgagattgt acaactgcga 3060
agtgaagtgg accatctccg ccgggagatc acggaacgag agatgcagct taccagccag 3120
aagcaaacga tggaggctct gaagaccacg tgcaccatgc tggaggaaca ggtcatggat 3180
ttggaggccc taaacgatga gctgctagaa aaagagcggc agtgggaggc ctggaggagc 3240
gtcctgggtg atgagaaatc ccagtttgag tgtcgggttc gagagctgca gagaatgctg 3300
gacaccgaga aacagagcag ggcgagagcc gatcagcgga tcaccgagtc tcgccaggtg 3360
gtggagctgg cagtgaagga gcacaaggct gagattctcg ctctgcagca ggctctcaaa 3420
gagcagaagc tgaaggccga gagcctctct gacaagctca atgacctgga gaagaagcat 3480
gctatgcttg aaatgaatgc ccgaagctta cagcagaagc tggagactga acgagagctc 3540
aaacagaggc ttctggaaga gcaagccaaa ttacagcagc agatggacct gcagaaaaat 3600
cacattttcc gtctgactca aggactgcaa gaagctctag atcgggctga tctactgaag 3660
acagaaagaa gtgacttgga gtatcagctg gaaaacattc aggttctcta ttctcatgaa 3720
aaggtgaaaa tggaaggcac tatttctcaa caaaccaaac tcattgattt tctgcaagcc 3780
aaaatggacc aacctgctaa aaagaaaaag ggtttattta gtcgacggaa agaggaccct 3840
gctttaccca cacaggttcc tctgcagtac aatgagctga agctggccct ggagaaggag 3900
aaagctcgct gtgcagagct agaggaagcc cttcagaaga cccgcatcga gctccggtcc 3960
gcccgggagg aagctgccca ccgcaaagca acggaccacc cacacccatc cacgccagcc 4020
accgcgaggc agcagatcgc catgtccgcc atcgtgcggt cgccagagca ccagcccagt 4080
gccatgagcc tgctggcccc gccatccagc cgcagaaagg agtcttcaac tccagaggaa 4140
tttagtcggc gtcttaagga acgcatgcac cacaatattc ctcaccgatt caacgtagga 4200
ctgaacatgc gagccacaaa gtgtgctgtg tgtctggata ccgtgcactt tggacgccag 4260
gcatccaaat gtctcgaatg tcaggtgatg tgtcacccca agtgctccac gtgcttgcca 4320
gccacctgcg gcttgcctgc tgaatatgcc acacacttca ccgaggcctt ctgccgtgac 4380

```
aaaatgaact ccccaggtct ccagaccaag gagcccagca gcagcttgca cctggaaggg 4440
tggatgaagg tgcccaggaa taacaaacga ggacagcaag gctgggacag gaagtacatt 4500
gtcctggagg gatcaaaagt cctcatttat gacaatgaag ccagagaagc tggacagagg 4560
ccggtggaag aatttgagct gtgccttccc gacggggatg tatctattca tggtgccgtt 4620
ggtgcttccg aactcgcaaa tacagccaaa gcagaaaaag cagaagctga tgctaaactg 4680
cttggaaact ccctgctgaa actggaaggt gatgaccgtc tagacatgaa ctgcacgctg 4740
cccttcagtg accaggtggt gttggtgggc accgaggaag ggctctacgc cctgaatgtc 4800
ttgaaaaact ccctaaccca tgtcccagga attggagcag tcttccaaat ttatattatc 4860
aaggacctgg agaagctact catgatagca ggagaagagc gggcactgtg tcttgtggac 4920
gtgaagaaag tgaaacagtc cctggcccag tcccacctgc ctgcccagcc cgacatctca 4980
cccaacattt ttgaagctgt caagggctgc cacttgtttg gggcaggcaa gattgagaac 5040
gggctctgca tctgtgcagc catgcccagc aaagtcgtca ttctccgcta caacgaaaac 5100
ctcagcaaat actgcatccg gaaagagata gagacctcag agccctgcag ctgtatccac 5160
ttcaccaatt acagtatcct cattggaacc aataaattct acgaaatcga catgaagcag 5220
tacacgctcg aggaattcct ggataagaat gaccattcct tggcacctgc tgtgtttgcc 5280
gcctcttcca acagcttccc tgtctcaatc gtgcaggtga acagcgcagg gcagcgagag 5340
gagtacttgc tgtgtttcca cgaatttgga gtgttcgtgg attcttacgg aagacgtagc 5400
cgcacagacg atctcaagtg gagtcgctta cctttggcct ttgcctacag agaaccctat 5460
ctgtttgtga cccacttcaa ctcactcgaa gtaattgaga tccaggcacg ctcctcagca 5520
gggacccctg cccgagcgta cctggacatc ccgaacccgc gctacctggg ccctgccatt 5580
tcctcaggag cgatttactt ggcgtcctca taccaggata aattaagggt catttgctgc 5640
aagggaaacc tcgtgaagga gtccggcact gaacaccacc ggggcccgtc cacctcccgc 5700
agcagcccca acaagcgagg cccacccacg tacaacgagc acatcaccaa gcgcgtggcc 5760
tccagcccag cgccgcccga aggccccagc cacccgcgag agccaagcac accccaccgc 5820
taccgcgagg ggcggaccga gctgcgcagg gacaagtctc ctggccgccc cctggagcga 5880
gagaagtccc ccggccggat gctcagcacg cggagagagc ggtcccccgg gaggctgttt 5940
gaagacagca gcaggggccg gctgcctgcg ggagccgtga ggaccccgct gtcccaggtg 6000
aacaagggaa gagggcagag tgcctctcaa gttttcacgg ttaacactgt cacctattat 6060
gactggaata aaaagctgga caacctgcca gctaactggt cagtcctgag gatcatccag 6120
ctgaatggag aaatccggca gcaggttgaa aagtctgttc tgagaacaga ttattgctga 6180
gcagagttca tgtgacttct agacgtggtg acttaaaaaa tggccttaag gctgcagagc 6240
cagccacctc tgcttacaaa aagagtactt agtgcacatg actgtaagaa acaattgtaa 6300
aacctcatct agaaatcaga aagcttctaa tttctataga aatgacacct ccctggagcc 6360
gagagacaat ctgttgttga ttttgaagga caggcaagac caacactgta tttagttcca 6420
tagccaggcc tcaacaggga caagtggctg gccttaaaaa cacacagatg actggaaatg 6480
atgtgtggcc tcagtccctg tttcccagaa ttttactggc aaaggagtta gcattcattt 6540
ttggcttaag aaaaatcgag aatgtaggtt taga                           6574

<210> 11
<211> 6162
<212> DNA
<213> Homo Sapiens

<220>
<221> CDS
<222> (1)...(6162)

<400> 11
atg ttg aag ttc aaa tat gga gcg cgg aat cct ttg gat gct ggt gct   48
Met Leu Lys Phe Lys Tyr Gly Ala Arg Asn Pro Leu Asp Ala Gly Ala
 1               5                  10                  15

gct gaa ccc att gcc agc cgg gcc tcc agg ctg aat ctg ttc ttc cag   96
Ala Glu Pro Ile Ala Ser Arg Ala Ser Arg Leu Asn Leu Phe Phe Gln
            20                  25                  30

ggg aaa cca ccc ttt atg act caa cag cag atg tct cct ctt tcc cga   144
Gly Lys Pro Pro Phe Met Thr Gln Gln Gln Met Ser Pro Leu Ser Arg
        35                  40                  45

gaa ggg ata tta gat gcc ctc ttt gtt ctc ttt gaa gaa tgc agt cag   192
Glu Gly Ile Leu Asp Ala Leu Phe Val Leu Phe Glu Glu Cys Ser Gln
```

```
     50                  55                  60

cct gct ctg atg aag att aag cac gtg agc aac ttt gtc cgg aag tat   240
Pro Ala Leu Met Lys Ile Lys His Val Ser Asn Phe Val Arg Lys Tyr
 65                  70                  75                  80

tcc gac acc ata gct gag tta cag gag ctc cag cct tcg gca aag gac   288
Ser Asp Thr Ile Ala Glu Leu Gln Glu Leu Gln Pro Ser Ala Lys Asp
                 85                  90                  95

ttc gaa gtc aga agt ctt gta ggt tgt ggt cac ttt gct gaa gtg cag   336
Phe Glu Val Arg Ser Leu Val Gly Cys Gly His Phe Ala Glu Val Gln
            100                 105                 110

gtg gta aga gag aaa gca acc ggg gac atc tat gct atg aaa gtg atg   384
Val Val Arg Glu Lys Ala Thr Gly Asp Ile Tyr Ala Met Lys Val Met
        115                 120                 125

aag aag aag gct tta ttg gcc cag gag cag gtt tca ttt ttt gag gaa   432
Lys Lys Lys Ala Leu Leu Ala Gln Glu Gln Val Ser Phe Phe Glu Glu
    130                 135                 140

gag cgg aac ata tta tct cga agc aca agc ccg tgg atc ccc caa tta   480
Glu Arg Asn Ile Leu Ser Arg Ser Thr Ser Pro Trp Ile Pro Gln Leu
145                 150                 155                 160

cag tat gcc ttt cag gac aaa aat cac ctt tat ctg atg gag gaa tat   528
Gln Tyr Ala Phe Gln Asp Lys Asn His Leu Tyr Leu Met Glu Glu Tyr
                165                 170                 175

cag cct gga ggg gac ttg ctg tca ctt ttg aat aga tat gag gac cag   576
Gln Pro Gly Gly Asp Leu Leu Ser Leu Leu Asn Arg Tyr Glu Asp Gln
            180                 185                 190

tta gat gaa aac ctg ata cag ttt tac cta gct gag ctg att ttg gct   624
Leu Asp Glu Asn Leu Ile Gln Phe Tyr Leu Ala Glu Leu Ile Leu Ala
        195                 200                 205

gtt cac agc gtt cat ctg atg gga tac gtg cat cga gac atc aag cct   672
Val His Ser Val His Leu Met Gly Tyr Val His Arg Asp Ile Lys Pro
    210                 215                 220

gag aac att ctc gtt gac cgc aca gga cac atc aag ctg gtg gat ttt   720
Glu Asn Ile Leu Val Asp Arg Thr Gly His Ile Lys Leu Val Asp Phe
225                 230                 235                 240

gga tct gcc gcg aaa atg aat tca aac aag atg gtg aat gcc aaa ctc   768
Gly Ser Ala Ala Lys Met Asn Ser Asn Lys Met Val Asn Ala Lys Leu
                245                 250                 255

ccg att ggg acc cca gat tac atg gct cct gaa gtg ctg act gtg atg   816
Pro Ile Gly Thr Pro Asp Tyr Met Ala Pro Glu Val Leu Thr Val Met
            260                 265                 270

aac ggg gat gga aaa ggc acc tac ggc ctg gac tgt gac tgg tgg tca   864
Asn Gly Asp Gly Lys Gly Thr Tyr Gly Leu Asp Cys Asp Trp Trp Ser
        275                 280                 285

gtg ggc gtg att gcc tat gag atg att tat ggg aga tcc ccc ttc gca   912
Val Gly Val Ile Ala Tyr Glu Met Ile Tyr Gly Arg Ser Pro Phe Ala
    290                 295                 300
```

```
gag gga acc tct gcc aga acc ttc aat aac att atg aat ttc cag cgg   960
Glu Gly Thr Ser Ala Arg Thr Phe Asn Asn Ile Met Asn Phe Gln Arg
305                 310                 315                 320

ttt ttg aaa ttt cca gat gac ccc aaa gtg agc agt gac ttt ctt gat   1008
Phe Leu Lys Phe Pro Asp Asp Pro Lys Val Ser Ser Asp Phe Leu Asp
                325                 330                 335

ctg att caa agc ttg ttg tgc ggc cag aaa gag aga ctg aag ttt gaa   1056
Leu Ile Gln Ser Leu Leu Cys Gly Gln Lys Glu Arg Leu Lys Phe Glu
            340                 345                 350

ggt ctt tgc tgc cat cct ttc ttc tct aaa att gac tgg aac aac att   1104
Gly Leu Cys Cys His Pro Phe Phe Ser Lys Ile Asp Trp Asn Asn Ile
        355                 360                 365

cgt aac tct cct ccc ccc ttc gtt ccc acc ctc aag tct gac gat gac   1152
Arg Asn Ser Pro Pro Pro Phe Val Pro Thr Leu Lys Ser Asp Asp Asp
    370                 375                 380

acc tcc aat ttt gat gaa cca gag aag aat tcg tgg gtt tca tcc tct   1200
Thr Ser Asn Phe Asp Glu Pro Glu Lys Asn Ser Trp Val Ser Ser Ser
385                 390                 395                 400

ccg tgc cag ctg agc ccc tca ggc ttc tcg ggt gaa gaa ctg ccg ttt   1248
Pro Cys Gln Leu Ser Pro Ser Gly Phe Ser Gly Glu Glu Leu Pro Phe
                405                 410                 415

gtg ggg ttt tcg tac agc aag gca ctg ggg att ctt ggt aga tct gag   1296
Val Gly Phe Ser Tyr Ser Lys Ala Leu Gly Ile Leu Gly Arg Ser Glu
            420                 425                 430

tct gtt gtg tcg ggt ctg gac tcc cct gcc aag act agc tcc atg gaa   1344
Ser Val Val Ser Gly Leu Asp Ser Pro Ala Lys Thr Ser Ser Met Glu
        435                 440                 445

aag aaa ctt ctc atc aaa agc aaa gag cta caa gac tct cag gac aag   1392
Lys Lys Leu Leu Ile Lys Ser Lys Glu Leu Gln Asp Ser Gln Asp Lys
    450                 455                 460

tgt cac aag atg gag cag gaa atg acc cgg tta cat cgg aga gtg tca   1440
Cys His Lys Met Glu Gln Glu Met Thr Arg Leu His Arg Arg Val Ser
465                 470                 475                 480

gag gtg gag gct gtg ctt agt cag aag gag gtg gag ctg aag gcc tct   1488
Glu Val Glu Ala Val Leu Ser Gln Lys Glu Val Glu Leu Lys Ala Ser
                485                 490                 495

gag act cag aga tcc ctc ctg gag cag gac ctt gct acc tac atc aca   1536
Glu Thr Gln Arg Ser Leu Leu Glu Gln Asp Leu Ala Thr Tyr Ile Thr
            500                 505                 510

gaa tgc agt agc tta aag cga agt ttg gag caa gca cgg atg gag gtg   1584
Glu Cys Ser Ser Leu Lys Arg Ser Leu Glu Gln Ala Arg Met Glu Val
        515                 520                 525

tcc cag gag gat gac aaa gca ctg cag ctt ctc cat gat atc aga gag   1632
Ser Gln Glu Asp Asp Lys Ala Leu Gln Leu Leu His Asp Ile Arg Glu
    530                 535                 540

cag agc cgg aag ctc caa gaa atc aaa gag cag gag tac cag gct caa   1680
Gln Ser Arg Lys Leu Gln Glu Ile Lys Glu Gln Glu Tyr Gln Ala Gln
```

```
545                 550                 555                 560

gtg gaa gaa atg agg ttg atg atg aat cag ttg gaa gag gat ctt gtc   1728
Val Glu Glu Met Arg Leu Met Met Asn Gln Leu Glu Glu Asp Leu Val
                565                 570                 575

tca gca aga aga cgg agt gat ctc tac gaa tct gag ctg aga gag tct   1776
Ser Ala Arg Arg Arg Ser Asp Leu Tyr Glu Ser Glu Leu Arg Glu Ser
            580                 585                 590

cgg ctt gct gct gaa gaa ttc aag cgg aaa gcg aca gaa tgt cag cat   1824
Arg Leu Ala Ala Glu Glu Phe Lys Arg Lys Ala Thr Glu Cys Gln His
        595                 600                 605

aaa ctg ttg aag gct aag gat caa ggg aag cct gaa gtg gga gaa tat   1872
Lys Leu Leu Lys Ala Lys Asp Gln Gly Lys Pro Glu Val Gly Glu Tyr
    610                 615                 620

gcg aaa ctg gag aag atc aat gct gag cag cag ctc aaa att cag gag   1920
Ala Lys Leu Glu Lys Ile Asn Ala Glu Gln Gln Leu Lys Ile Gln Glu
625                 630                 635                 640

ctc caa gag aaa ctg gag aag gct gca aag gag cga gcc gag agg gag   1968
Leu Gln Glu Lys Leu Glu Lys Ala Ala Lys Glu Arg Ala Glu Arg Glu
                645                 650                 655

ctg gag aag ctg cag aac cga gag gat tct tct gaa ggc atc aga aag   2016
Leu Glu Lys Leu Gln Asn Arg Glu Asp Ser Ser Glu Gly Ile Arg Lys
            660                 665                 670

aag ctg gtg gaa gct gag gaa cgc cgc cat tct ctg gag aac aag gta   2064
Lys Leu Val Glu Ala Glu Glu Arg Arg His Ser Leu Glu Asn Lys Val
        675                 680                 685

aag aga cta gag acc atg gag cgt aga gaa aac aga ctg aag gat gac   2112
Lys Arg Leu Glu Thr Met Glu Arg Arg Glu Asn Arg Leu Lys Asp Asp
    690                 695                 700

atc cag aca aaa tcc caa cag atc cag cag atg gct gat aaa att ctg   2160
Ile Gln Thr Lys Ser Gln Gln Ile Gln Gln Met Ala Asp Lys Ile Leu
705                 710                 715                 720

gag ctc gaa gag aaa cat cgg gag gcc caa gtc tca gcc cag cac cta   2208
Glu Leu Glu Glu Lys His Arg Glu Ala Gln Val Ser Ala Gln His Leu
                725                 730                 735

gaa gtg cac ctg aaa cag aaa gag cag cac tat gag gaa aag att aaa   2256
Glu Val His Leu Lys Gln Lys Glu Gln His Tyr Glu Glu Lys Ile Lys
            740                 745                 750

gtg ttg gac aat cag ata aag aaa gac ctg gct gac aag gag aca ctg   2304
Val Leu Asp Asn Gln Ile Lys Lys Asp Leu Ala Asp Lys Glu Thr Leu
        755                 760                 765

gag aac atg atg cag aga cac gag gag gag gcc cat gag aag ggc aaa   2352
Glu Asn Met Met Gln Arg His Glu Glu Glu Ala His Glu Lys Gly Lys
    770                 775                 780

att ctc agc gaa cag aag gcg atg atc aat gct atg gat tcc aag atc   2400
Ile Leu Ser Glu Gln Lys Ala Met Ile Asn Ala Met Asp Ser Lys Ile
785                 790                 795                 800
```

```
aga tcc ctg gaa cag agg att gtg gaa ctg tct gaa gcc aat aaa ctt   2448
Arg Ser Leu Glu Gln Arg Ile Val Glu Leu Ser Glu Ala Asn Lys Leu
                805                 810                 815

gca gca aat agc agt ctt ttt acc caa agg aac atg aag gcc caa gaa   2496
Ala Ala Asn Ser Ser Leu Phe Thr Gln Arg Asn Met Lys Ala Gln Glu
            820                 825                 830

gag atg att tct gaa ctc agg caa cag aaa ttt tac ctg gag aca cag   2544
Glu Met Ile Ser Glu Leu Arg Gln Gln Lys Phe Tyr Leu Glu Thr Gln
        835                 840                 845

gct ggg aag ttg gag gcc cag aac cga aaa ctg gag gag cag ctg gag   2592
Ala Gly Lys Leu Glu Ala Gln Asn Arg Lys Leu Glu Glu Gln Leu Glu
    850                 855                 860

aag atc agc cac caa gac cac agt gac aag aat cgg ctg ctg gaa ctg   2640
Lys Ile Ser His Gln Asp His Ser Asp Lys Asn Arg Leu Leu Glu Leu
865                 870                 875                 880

gag aca aga ttg cgg gag gtc agt cta gag cac gag gag cag aaa ctg   2688
Glu Thr Arg Leu Arg Glu Val Ser Leu Glu His Glu Glu Gln Lys Leu
                885                 890                 895

gag ctc aag cgc cag ctc aca gag cta cag ctc tcc ctg cag gag cgc   2736
Glu Leu Lys Arg Gln Leu Thr Glu Leu Gln Leu Ser Leu Gln Glu Arg
            900                 905                 910

gag tca cag ttg aca gcc ctg cag gct gca cgg gcg gcc ctg gag agc   2784
Glu Ser Gln Leu Thr Ala Leu Gln Ala Ala Arg Ala Ala Leu Glu Ser
        915                 920                 925

cag ctt cgc cag gcg aag aca gag ctg gaa gag acc aca gca gaa gct   2832
Gln Leu Arg Gln Ala Lys Thr Glu Leu Glu Glu Thr Thr Ala Glu Ala
    930                 935                 940

gaa gag gag atc cag gca ctc acg gca cat aga gat gaa atc cag cgc   2880
Glu Glu Glu Ile Gln Ala Leu Thr Ala His Arg Asp Glu Ile Gln Arg
945                 950                 955                 960

aaa ttt gat gct ctt cgt aac agc tgt act gta atc aca gac ctg gag   2928
Lys Phe Asp Ala Leu Arg Asn Ser Cys Thr Val Ile Thr Asp Leu Glu
                965                 970                 975

gag cag cta aac cag ctg acc gag gac aac gct gaa ctc aac aac caa   2976
Glu Gln Leu Asn Gln Leu Thr Glu Asp Asn Ala Glu Leu Asn Asn Gln
            980                 985                 990

aac ttc tac ttg tcc aaa caa ctc gat gag gct tct ggc gcc aac gac   3024
Asn Phe Tyr Leu Ser Lys Gln Leu Asp Glu Ala Ser Gly Ala Asn Asp
        995                 1000                1005

gag att gta caa ctg cga agt gaa gtg gac cat ctc cgc cgg gag atc   3072
Glu Ile Val Gln Leu Arg Ser Glu Val Asp His Leu Arg Arg Glu Ile
    1010                1015                1020

acg gaa cga gag atg cag ctt acc agc cag aag caa acg atg gag gct   3120
Thr Glu Arg Glu Met Gln Leu Thr Ser Gln Lys Gln Thr Met Glu Ala
1025                1030                1035                1040

ctg aag acc acg tgc acc atg ctg gag gaa cag gtc atg gat ttg gag   3168
Leu Lys Thr Thr Cys Thr Met Leu Glu Glu Gln Val Met Asp Leu Glu
```

```
                1045                1050                1055

gcc cta aac gat gag ctg cta gaa aaa gag cgg cag tgg gag gcc tgg   3216
Ala Leu Asn Asp Glu Leu Leu Glu Lys Glu Arg Gln Trp Glu Ala Trp
            1060                1065                1070

agg agc gtc ctg ggt gat gag aaa tcc cag ttt gag tgt cgg gtt cga   3264
Arg Ser Val Leu Gly Asp Glu Lys Ser Gln Phe Glu Cys Arg Val Arg
        1075                1080                1085

gag ctg cag aga atg ctg gac acc gag aaa cag agc agg gcg aga gcc   3312
Glu Leu Gln Arg Met Leu Asp Thr Glu Lys Gln Ser Arg Ala Arg Ala
    1090                1095                1100

gat cag cgg atc acc gag tct cgc cag gtg gtg gag ctg gca gtg aag   3360
Asp Gln Arg Ile Thr Glu Ser Arg Gln Val Val Glu Leu Ala Val Lys
1105                1110                1115                1120

gag cac aag gct gag att ctc gct ctg cag cag gct ctc aaa gag cag   3408
Glu His Lys Ala Glu Ile Leu Ala Leu Gln Gln Ala Leu Lys Glu Gln
                1125                1130                1135

aag ctg aag gcc gag agc ctc tct gac aag ctc aat gac ctg gag aag   3456
Lys Leu Lys Ala Glu Ser Leu Ser Asp Lys Leu Asn Asp Leu Glu Lys
            1140                1145                1150

aag cat gct atg ctt gaa atg aat gcc cga agc tta cag cag aag ctg   3504
Lys His Ala Met Leu Glu Met Asn Ala Arg Ser Leu Gln Gln Lys Leu
        1155                1160                1165

gag act gaa cga gag ctc aaa cag agg ctt ctg gaa gag caa gcc aaa   3552
Glu Thr Glu Arg Glu Leu Lys Gln Arg Leu Leu Glu Glu Gln Ala Lys
    1170                1175                1180

tta cag cag cag atg gac ctg cag aaa aat cac att ttc cgt ctg act   3600
Leu Gln Gln Gln Met Asp Leu Gln Lys Asn His Ile Phe Arg Leu Thr
1185                1190                1195                1200

caa gga ctg caa gaa gct cta gat cgg gct gat cta ctg aag aca gaa   3648
Gln Gly Leu Gln Glu Ala Leu Asp Arg Ala Asp Leu Leu Lys Thr Glu
                1205                1210                1215

aga agt gac ttg gag tat cag ctg gaa aac att cag gtt ctc tat tct   3696
Arg Ser Asp Leu Glu Tyr Gln Leu Glu Asn Ile Gln Val Leu Tyr Ser
            1220                1225                1230

cat gaa aag gtg aaa atg gaa ggc act att tct caa caa acc aaa ctc   3744
His Glu Lys Val Lys Met Glu Gly Thr Ile Ser Gln Gln Thr Lys Leu
        1235                1240                1245

att gat ttt ctg caa gcc aaa atg gac caa cct gct aaa aag aaa aag   3792
Ile Asp Phe Leu Gln Ala Lys Met Asp Gln Pro Ala Lys Lys Lys Lys
    1250                1255                1260

ggt tta ttt agt cga cgg aaa gag gac cct gct tta ccc aca cag gtt   3840
Gly Leu Phe Ser Arg Arg Lys Glu Asp Pro Ala Leu Pro Thr Gln Val
1265                1270                1275                1280

cct ctg cag tac aat gag ctg aag ctg gcc ctg gag aag gag aaa gct   3888
Pro Leu Gln Tyr Asn Glu Leu Lys Leu Ala Leu Glu Lys Glu Lys Ala
                1285                1290                1295
```

```
cgc tgt gca gag cta gag gaa gcc ctt cag aag acc cgc atc gag ctc   3936
Arg Cys Ala Glu Leu Glu Glu Ala Leu Gln Lys Thr Arg Ile Glu Leu
            1300                1305                1310

cgg tcc gcc cgg gag gaa gct gcc cac cgc aaa gca acg gac cac cca   3984
Arg Ser Ala Arg Glu Glu Ala Ala His Arg Lys Ala Thr Asp His Pro
        1315                1320                1325

cac cca tcc acg cca gcc acc gcg agg cag cag atc gcc atg tcc gcc   4032
His Pro Ser Thr Pro Ala Thr Ala Arg Gln Gln Ile Ala Met Ser Ala
    1330                1335                1340

atc gtg cgg tcg cca gag cac cag ccc agt gcc atg agc ctg ctg gcc   4080
Ile Val Arg Ser Pro Glu His Gln Pro Ser Ala Met Ser Leu Leu Ala
1345                1350                1355                1360

ccg cca tcc agc cgc aga aag gag tct tca act cca gag gaa ttt agt   4128
Pro Pro Ser Ser Arg Arg Lys Glu Ser Ser Thr Pro Glu Glu Phe Ser
                1365                1370                1375

cgg cgt ctt aag gaa cgc atg cac cac aat att cct cac cga ttc aac   4176
Arg Arg Leu Lys Glu Arg Met His His Asn Ile Pro His Arg Phe Asn
            1380                1385                1390

gta gga ctg aac atg cga gcc aca aag tgt gct gtg tgt ctg gat acc   4224
Val Gly Leu Asn Met Arg Ala Thr Lys Cys Ala Val Cys Leu Asp Thr
        1395                1400                1405

gtg cac ttt gga cgc cag gca tcc aaa tgt ctc gaa tgt cag gtg atg   4272
Val His Phe Gly Arg Gln Ala Ser Lys Cys Leu Glu Cys Gln Val Met
    1410                1415                1420

tgt cac ccc aag tgc tcc acg tgc ttg cca gcc acc tgc ggc ttg cct   4320
Cys His Pro Lys Cys Ser Thr Cys Leu Pro Ala Thr Cys Gly Leu Pro
1425                1430                1435                1440

gct gaa tat gcc aca cac ttc acc gag gcc ttc tgc cgt gac aaa atg   4368
Ala Glu Tyr Ala Thr His Phe Thr Glu Ala Phe Cys Arg Asp Lys Met
                1445                1450                1455

aac tcc cca ggt ctc cag acc aag gag ccc agc agc agc ttg cac ctg   4416
Asn Ser Pro Gly Leu Gln Thr Lys Glu Pro Ser Ser Ser Leu His Leu
            1460                1465                1470

gaa ggg tgg atg aag gtg ccc agg aat aac aaa cga gga cag caa ggc   4464
Glu Gly Trp Met Lys Val Pro Arg Asn Asn Lys Arg Gly Gln Gln Gly
        1475                1480                1485

tgg gac agg aag tac att gtc ctg gag gga tca aaa gtc ctc att tat   4512
Trp Asp Arg Lys Tyr Ile Val Leu Glu Gly Ser Lys Val Leu Ile Tyr
    1490                1495                1500

gac aat gaa gcc aga gaa gct gga cag agg ccg gtg gaa gaa ttt gag   4560
Asp Asn Glu Ala Arg Glu Ala Gly Gln Arg Pro Val Glu Glu Phe Glu
1505                1510                1515                1520

ctg tgc ctt ccc gac ggg gat gta tct att cat ggt gcc gtt ggt gct   4608
Leu Cys Leu Pro Asp Gly Asp Val Ser Ile His Gly Ala Val Gly Ala
                1525                1530                1535

tcc gaa ctc gca aat aca gcc aaa gca gaa aaa gca gaa gct gat gct   4656
Ser Glu Leu Ala Asn Thr Ala Lys Ala Glu Lys Ala Glu Ala Asp Ala
```

```
            1540                1545                1550

aaa ctg ctt gga aac tcc ctg ctg aaa ctg gaa ggt gat gac cgt cta   4704
Lys Leu Leu Gly Asn Ser Leu Leu Lys Leu Glu Gly Asp Asp Arg Leu
        1555                1560                1565

gac atg aac tgc acg ctg ccc ttc agt gac cag gtg gtg ttg gtg ggc   4752
Asp Met Asn Cys Thr Leu Pro Phe Ser Asp Gln Val Val Leu Val Gly
    1570                1575                1580

acc gag gaa ggg ctc tac gcc ctg aat gtc ttg aaa aac tcc cta acc   4800
Thr Glu Glu Gly Leu Tyr Ala Leu Asn Val Leu Lys Asn Ser Leu Thr
1585                1590                1595                1600

cat gtc cca gga att gga gca gtc ttc caa att tat att atc aag gac   4848
His Val Pro Gly Ile Gly Ala Val Phe Gln Ile Tyr Ile Ile Lys Asp
                1605                1610                1615

ctg gag aag cta ctc atg ata gca gga gaa gag cgg gca ctg tgt ctt   4896
Leu Glu Lys Leu Leu Met Ile Ala Gly Glu Glu Arg Ala Leu Cys Leu
            1620                1625                1630

gtg gac gtg aag aaa gtg aaa cag tcc ctg gcc cag tcc cac ctg cct   4944
Val Asp Val Lys Lys Val Lys Gln Ser Leu Ala Gln Ser His Leu Pro
        1635                1640                1645

gcc cag ccc gac atc tca ccc aac att ttt gaa gct gtc aag ggc tgc   4992
Ala Gln Pro Asp Ile Ser Pro Asn Ile Phe Glu Ala Val Lys Gly Cys
    1650                1655                1660

cac ttg ttt ggg gca ggc aag att gag aac ggg ctc tgc atc tgt gca   5040
His Leu Phe Gly Ala Gly Lys Ile Glu Asn Gly Leu Cys Ile Cys Ala
1665                1670                1675                1680

gcc atg ccc agc aaa gtc gtc att ctc cgc tac aac gaa aac ctc agc   5088
Ala Met Pro Ser Lys Val Val Ile Leu Arg Tyr Asn Glu Asn Leu Ser
                1685                1690                1695

aaa tac tgc atc cgg aaa gag ata gag acc tca gag ccc tgc agc tgt   5136
Lys Tyr Cys Ile Arg Lys Glu Ile Glu Thr Ser Glu Pro Cys Ser Cys
            1700                1705                1710

atc cac ttc acc aat tac agt atc ctc att gga acc aat aaa ttc tac   5184
Ile His Phe Thr Asn Tyr Ser Ile Leu Ile Gly Thr Asn Lys Phe Tyr
        1715                1720                1725

gaa atc gac atg aag cag tac acg ctc gag gaa ttc ctg gat aag aat   5232
Glu Ile Asp Met Lys Gln Tyr Thr Leu Glu Glu Phe Leu Asp Lys Asn
    1730                1735                1740

gac cat tcc ttg gca cct gct gtg ttt gcc gcc tct tcc aac agc ttc   5280
Asp His Ser Leu Ala Pro Ala Val Phe Ala Ala Ser Ser Asn Ser Phe
1745                1750                1755                1760

cct gtc tca atc gtg cag gtg aac agc gca ggg cag cga gag gag tac   5328
Pro Val Ser Ile Val Gln Val Asn Ser Ala Gly Gln Arg Glu Glu Tyr
                1765                1770                1775

ttg ctg tgt ttc cac gaa ttt gga gtg ttc gtg gat tct tac gga aga   5376
Leu Leu Cys Phe His Glu Phe Gly Val Phe Val Asp Ser Tyr Gly Arg
            1780                1785                1790
```

```
cgt agc cgc aca gac gat ctc aag tgg agt cgc tta cct ttg gcc ttt   5424
Arg Ser Arg Thr Asp Asp Leu Lys Trp Ser Arg Leu Pro Leu Ala Phe
        1795                1800                1805

gcc tac aga gaa ccc tat ctg ttt gtg acc cac ttc aac tca ctc gaa   5472
Ala Tyr Arg Glu Pro Tyr Leu Phe Val Thr His Phe Asn Ser Leu Glu
    1810                1815                1820

gta att gag atc cag gca cgc tcc tca gca ggg acc cct gcc cga gcg   5520
Val Ile Glu Ile Gln Ala Arg Ser Ser Ala Gly Thr Pro Ala Arg Ala
1825                1830                1835                1840

tac ctg gac atc ccg aac ccg cgc tac ctg ggc cct gcc att tcc tca   5568
Tyr Leu Asp Ile Pro Asn Pro Arg Tyr Leu Gly Pro Ala Ile Ser Ser
                1845                1850                1855

gga gcg att tac ttg gcg tcc tca tac cag gat aaa tta agg gtc att   5616
Gly Ala Ile Tyr Leu Ala Ser Ser Tyr Gln Asp Lys Leu Arg Val Ile
            1860                1865                1870

tgc tgc aag gga aac ctc gtg aag gag tcc ggc act gaa cac cac cgg   5664
Cys Cys Lys Gly Asn Leu Val Lys Glu Ser Gly Thr Glu His His Arg
        1875                1880                1885

ggc ccg tcc acc tcc cgc agc agc ccc aac aag cga ggc cca ccc acg   5712
Gly Pro Ser Thr Ser Arg Ser Ser Pro Asn Lys Arg Gly Pro Pro Thr
    1890                1895                1900

tac aac gag cac atc acc aag cgc gtg gcc tcc agc cca gcg ccg ccc   5760
Tyr Asn Glu His Ile Thr Lys Arg Val Ala Ser Ser Pro Ala Pro Pro
1905                1910                1915                1920

gaa ggc ccc agc cac ccg cga gag cca agc aca ccc cac cgc tac cgc   5808
Glu Gly Pro Ser His Pro Arg Glu Pro Ser Thr Pro His Arg Tyr Arg
                1925                1930                1935

gag ggg cgg acc gag ctg cgc agg gac aag tct cct ggc cgc ccc ctg   5856
Glu Gly Arg Thr Glu Leu Arg Arg Asp Lys Ser Pro Gly Arg Pro Leu
            1940                1945                1950

gag cga gag aag tcc ccc ggc cgg atg ctc agc acg cgg aga gag cgg   5904
Glu Arg Glu Lys Ser Pro Gly Arg Met Leu Ser Thr Arg Arg Glu Arg
        1955                1960                1965

tcc ccc ggg agg ctg ttt gaa gac agc agc agg ggc cgg ctg cct gcg   5952
Ser Pro Gly Arg Leu Phe Glu Asp Ser Ser Arg Gly Arg Leu Pro Ala
    1970                1975                1980

gga gcc gtg agg acc ccg ctg tcc cag gtg aac aag gga aga ggg cag   6000
Gly Ala Val Arg Thr Pro Leu Ser Gln Val Asn Lys Gly Arg Gly Gln
1985                1990                1995                2000

agt gcc tct caa gtt ttc acg gtt aac act gtc acc tat tat gac tgg   6048
Ser Ala Ser Gln Val Phe Thr Val Asn Thr Val Thr Tyr Tyr Asp Trp
                2005                2010                2015

aat aaa aag ctg gac aac ctg cca gct aac tgg tca gtc ctg agg atc   6096
Asn Lys Lys Leu Asp Asn Leu Pro Ala Asn Trp Ser Val Leu Arg Ile
            2020                2025                2030

atc cag ctg aat gga gaa atc cgg cag cag gtt gaa aag tct gtt ctg   6144
Ile Gln Leu Asn Gly Glu Ile Arg Gln Gln Val Glu Lys Ser Val Leu
```

```
        2035                2040                2045

aga aca gat tat tgc tga                                                 6162
Arg Thr Asp Tyr Cys  *
    2050


<210> 12
<211> 2053
<212> PRT
<213> Homo Sapiens

<400> 12
Met Leu Lys Phe Lys Tyr Gly Ala Arg Asn Pro Leu Asp Ala Gly Ala
 1               5                  10                  15
Ala Glu Pro Ile Ala Ser Arg Ala Ser Arg Leu Asn Leu Phe Phe Gln
            20                  25                  30
Gly Lys Pro Pro Phe Met Thr Gln Gln Gln Met Ser Pro Leu Ser Arg
        35                  40                  45
Glu Gly Ile Leu Asp Ala Leu Phe Val Leu Phe Glu Glu Cys Ser Gln
    50                  55                  60
Pro Ala Leu Met Lys Ile Lys His Val Ser Asn Phe Val Arg Lys Tyr
65                  70                  75                  80
Ser Asp Thr Ile Ala Glu Leu Gln Glu Leu Gln Pro Ser Ala Lys Asp
                85                  90                  95
Phe Glu Val Arg Ser Leu Val Gly Cys Gly His Phe Ala Glu Val Gln
            100                 105                 110
Val Val Arg Glu Lys Ala Thr Gly Asp Ile Tyr Ala Met Lys Val Met
        115                 120                 125
Lys Lys Lys Ala Leu Leu Ala Gln Glu Gln Val Ser Phe Phe Glu Glu
    130                 135                 140
Glu Arg Asn Ile Leu Ser Arg Ser Thr Ser Pro Trp Ile Pro Gln Leu
145                 150                 155                 160
Gln Tyr Ala Phe Gln Asp Lys Asn His Leu Tyr Leu Met Glu Glu Tyr
                165                 170                 175
Gln Pro Gly Gly Asp Leu Leu Ser Leu Leu Asn Arg Tyr Glu Asp Gln
            180                 185                 190
Leu Asp Glu Asn Leu Ile Gln Phe Tyr Leu Ala Glu Leu Ile Leu Ala
        195                 200                 205
Val His Ser Val His Leu Met Gly Tyr Val His Arg Asp Ile Lys Pro
    210                 215                 220
Glu Asn Ile Leu Val Asp Arg Thr Gly His Ile Lys Leu Val Asp Phe
225                 230                 235                 240
Gly Ser Ala Ala Lys Met Asn Ser Asn Lys Met Val Asn Ala Lys Leu
                245                 250                 255
Pro Ile Gly Thr Pro Asp Tyr Met Ala Pro Glu Val Leu Thr Val Met
            260                 265                 270
Asn Gly Asp Gly Lys Gly Thr Tyr Gly Leu Asp Cys Asp Trp Trp Ser
        275                 280                 285
Val Gly Val Ile Ala Tyr Glu Met Ile Tyr Gly Arg Ser Pro Phe Ala
    290                 295                 300
Glu Gly Thr Ser Ala Arg Thr Phe Asn Asn Ile Met Asn Phe Gln Arg
305                 310                 315                 320
Phe Leu Lys Phe Pro Asp Asp Pro Lys Val Ser Ser Asp Phe Leu Asp
                325                 330                 335
Leu Ile Gln Ser Leu Leu Cys Gly Gln Lys Glu Arg Leu Lys Phe Glu
            340                 345                 350
Gly Leu Cys Cys His Pro Phe Phe Ser Lys Ile Asp Trp Asn Asn Ile
        355                 360                 365
Arg Asn Ser Pro Pro Pro Phe Val Pro Thr Leu Lys Ser Asp Asp Asp
    370                 375                 380
Thr Ser Asn Phe Asp Glu Pro Glu Lys Asn Ser Trp Val Ser Ser Ser
```

385 390 395 400
Pro Cys Gln Leu Ser Pro Ser Gly Phe Ser Gly Glu Glu Leu Pro Phe
405 410 415
Val Gly Phe Ser Tyr Ser Lys Ala Leu Gly Ile Leu Gly Arg Ser Glu
420 425 430
Ser Val Val Ser Gly Leu Asp Ser Pro Ala Lys Thr Ser Ser Met Glu
435 440 445
Lys Lys Leu Leu Ile Lys Ser Lys Glu Leu Gln Asp Ser Gln Asp Lys
450 455 460
Cys His Lys Met Glu Gln Glu Met Thr Arg Leu His Arg Arg Val Ser
465 470 475 480
Glu Val Glu Ala Val Leu Ser Gln Lys Glu Val Glu Leu Lys Ala Ser
485 490 495
Glu Thr Gln Arg Ser Leu Leu Glu Gln Asp Leu Ala Thr Tyr Ile Thr
500 505 510
Glu Cys Ser Ser Leu Lys Arg Ser Leu Glu Gln Ala Arg Met Glu Val
515 520 525
Ser Gln Glu Asp Asp Lys Ala Leu Gln Leu Leu His Asp Ile Arg Glu
530 535 540
Gln Ser Arg Lys Leu Gln Glu Ile Lys Glu Gln Glu Tyr Gln Ala Gln
545 550 555 560
Val Glu Glu Met Arg Leu Met Met Asn Gln Leu Glu Glu Asp Leu Val
565 570 575
Ser Ala Arg Arg Arg Ser Asp Leu Tyr Glu Ser Glu Leu Arg Glu Ser
580 585 590
Arg Leu Ala Ala Glu Glu Phe Lys Arg Lys Ala Thr Glu Cys Gln His
595 600 605
Lys Leu Leu Lys Ala Lys Asp Gln Gly Lys Pro Glu Val Gly Glu Tyr
610 615 620
Ala Lys Leu Glu Lys Ile Asn Ala Glu Gln Gln Leu Lys Ile Gln Glu
625 630 635 640
Leu Gln Glu Lys Leu Glu Lys Ala Ala Lys Glu Arg Ala Glu Arg Glu
645 650 655
Leu Glu Lys Leu Gln Asn Arg Glu Asp Ser Ser Glu Gly Ile Arg Lys
660 665 670
Lys Leu Val Glu Ala Glu Glu Arg Arg His Ser Leu Glu Asn Lys Val
675 680 685
Lys Arg Leu Glu Thr Met Glu Arg Arg Glu Asn Arg Leu Lys Asp Asp
690 695 700
Ile Gln Thr Lys Ser Gln Gln Ile Gln Gln Met Ala Asp Lys Ile Leu
705 710 715 720
Glu Leu Glu Glu Lys His Arg Glu Ala Gln Val Ser Ala Gln His Leu
725 730 735
Glu Val His Leu Lys Gln Lys Glu Gln His Tyr Glu Glu Lys Ile Lys
740 745 750
Val Leu Asp Asn Gln Ile Lys Lys Asp Leu Ala Asp Lys Glu Thr Leu
755 760 765
Glu Asn Met Met Gln Arg His Glu Glu Glu Ala His Glu Lys Gly Lys
770 775 780
Ile Leu Ser Glu Gln Lys Ala Met Ile Asn Ala Met Asp Ser Lys Ile
785 790 795 800
Arg Ser Leu Glu Gln Arg Ile Val Glu Leu Ser Glu Ala Asn Lys Leu
805 810 815
Ala Ala Asn Ser Ser Leu Phe Thr Gln Arg Asn Met Lys Ala Gln Glu
820 825 830
Glu Met Ile Ser Glu Leu Arg Gln Gln Lys Phe Tyr Leu Glu Thr Gln
835 840 845
Ala Gly Lys Leu Glu Ala Gln Asn Arg Lys Leu Glu Glu Gln Leu Glu
850 855 860
Lys Ile Ser His Gln Asp His Ser Asp Lys Asn Arg Leu Leu Glu Leu
865 870 875 880
Glu Thr Arg Leu Arg Glu Val Ser Leu Glu His Glu Glu Gln Lys Leu

```
                885                 890                 895
Glu Leu Lys Arg Gln Leu Thr Glu Leu Gln Leu Ser Leu Gln Glu Arg
            900                 905                 910
Glu Ser Gln Leu Thr Ala Leu Gln Ala Ala Arg Ala Ala Leu Glu Ser
        915                 920                 925
Gln Leu Arg Gln Ala Lys Thr Glu Leu Glu Glu Thr Thr Ala Glu Ala
    930                 935                 940
Glu Glu Glu Ile Gln Ala Leu Thr Ala His Arg Asp Glu Ile Gln Arg
945                 950                 955                 960
Lys Phe Asp Ala Leu Arg Asn Ser Cys Thr Val Ile Thr Asp Leu Glu
                965                 970                 975
Glu Gln Leu Asn Gln Leu Thr Glu Asp Asn Ala Glu Leu Asn Asn Gln
            980                 985                 990
Asn Phe Tyr Leu Ser Lys Gln Leu Asp Glu Ala Ser Gly Ala Asn Asp
        995                 1000                1005
Glu Ile Val Gln Leu Arg Ser Glu Val Asp His Leu Arg Arg Glu Ile
    1010                1015                1020
Thr Glu Arg Glu Met Gln Leu Thr Ser Gln Lys Gln Thr Met Glu Ala
1025                1030                1035                1040
Leu Lys Thr Thr Cys Thr Met Leu Glu Glu Gln Val Met Asp Leu Glu
                1045                1050                1055
Ala Leu Asn Asp Glu Leu Leu Glu Lys Glu Arg Gln Trp Glu Ala Trp
            1060                1065                1070
Arg Ser Val Leu Gly Asp Glu Lys Ser Gln Phe Glu Cys Arg Val Arg
        1075                1080                1085
Glu Leu Gln Arg Met Leu Asp Thr Glu Lys Gln Ser Arg Ala Arg Ala
    1090                1095                1100
Asp Gln Arg Ile Thr Glu Ser Arg Gln Val Val Glu Leu Ala Val Lys
1105                1110                1115                1120
Glu His Lys Ala Glu Ile Leu Ala Leu Gln Gln Ala Leu Lys Glu Gln
                1125                1130                1135
Lys Leu Lys Ala Glu Ser Leu Ser Asp Lys Leu Asn Asp Leu Glu Lys
            1140                1145                1150
Lys His Ala Met Leu Glu Met Asn Ala Arg Ser Leu Gln Gln Lys Leu
        1155                1160                1165
Glu Thr Glu Arg Glu Leu Lys Gln Arg Leu Leu Glu Glu Gln Ala Lys
    1170                1175                1180
Leu Gln Gln Gln Met Asp Leu Gln Lys Asn His Ile Phe Arg Leu Thr
1185                1190                1195                1200
Gln Gly Leu Gln Glu Ala Leu Asp Arg Ala Asp Leu Leu Lys Thr Glu
                1205                1210                1215
Arg Ser Asp Leu Glu Tyr Gln Leu Glu Asn Ile Gln Val Leu Tyr Ser
            1220                1225                1230
His Glu Lys Val Lys Met Glu Gly Thr Ile Ser Gln Gln Thr Lys Leu
        1235                1240                1245
Ile Asp Phe Leu Gln Ala Lys Met Asp Gln Pro Ala Lys Lys Lys Lys
    1250                1255                1260
Gly Leu Phe Ser Arg Arg Lys Glu Asp Pro Ala Leu Pro Thr Gln Val
1265                1270                1275                1280
Pro Leu Gln Tyr Asn Glu Leu Lys Leu Ala Leu Glu Lys Glu Lys Ala
                1285                1290                1295
Arg Cys Ala Glu Leu Glu Glu Ala Leu Gln Lys Thr Arg Ile Glu Leu
            1300                1305                1310
Arg Ser Ala Arg Glu Glu Ala Ala His Arg Lys Ala Thr Asp His Pro
        1315                1320                1325
His Pro Ser Thr Pro Ala Thr Ala Arg Gln Gln Ile Ala Met Ser Ala
    1330                1335                1340
Ile Val Arg Ser Pro Glu His Gln Pro Ser Ala Met Ser Leu Leu Ala
1345                1350                1355                1360
Pro Pro Ser Ser Arg Arg Lys Glu Ser Ser Thr Pro Glu Glu Phe Ser
                1365                1370                1375
Arg Arg Leu Lys Glu Arg Met His His Asn Ile Pro His Arg Phe Asn
```

```
            1380                1385                1390
Val Gly Leu Asn Met Arg Ala Thr Lys Cys Ala Val Cys Leu Asp Thr
        1395                1400                1405
Val His Phe Gly Arg Gln Ala Ser Lys Cys Leu Glu Cys Gln Val Met
    1410                1415                1420
Cys His Pro Lys Cys Ser Thr Cys Leu Pro Ala Thr Cys Gly Leu Pro
1425                1430                1435                1440
Ala Glu Tyr Ala Thr His Phe Thr Glu Ala Phe Cys Arg Asp Lys Met
                1445                1450                1455
Asn Ser Pro Gly Leu Gln Thr Lys Glu Pro Ser Ser Ser Leu His Leu
            1460                1465                1470
Glu Gly Trp Met Lys Val Pro Arg Asn Asn Lys Arg Gly Gln Gln Gly
        1475                1480                1485
Trp Asp Arg Lys Tyr Ile Val Leu Glu Gly Ser Lys Val Leu Ile Tyr
    1490                1495                1500
Asp Asn Glu Ala Arg Glu Ala Gly Gln Arg Pro Val Glu Glu Phe Glu
1505                1510                1515                1520
Leu Cys Leu Pro Asp Gly Asp Val Ser Ile His Gly Ala Val Gly Ala
                1525                1530                1535
Ser Glu Leu Ala Asn Thr Ala Lys Ala Glu Lys Ala Glu Ala Asp Ala
            1540                1545                1550
Lys Leu Leu Gly Asn Ser Leu Leu Lys Leu Glu Gly Asp Asp Arg Leu
        1555                1560                1565
Asp Met Asn Cys Thr Leu Pro Phe Ser Asp Gln Val Val Leu Val Gly
    1570                1575                1580
Thr Glu Glu Gly Leu Tyr Ala Leu Asn Val Leu Lys Asn Ser Leu Thr
1585                1590                1595                1600
His Val Pro Gly Ile Gly Ala Val Phe Gln Ile Tyr Ile Ile Lys Asp
                1605                1610                1615
Leu Glu Lys Leu Leu Met Ile Ala Gly Glu Glu Arg Ala Leu Cys Leu
            1620                1625                1630
Val Asp Val Lys Lys Val Lys Gln Ser Leu Ala Gln Ser His Leu Pro
        1635                1640                1645
Ala Gln Pro Asp Ile Ser Pro Asn Ile Phe Glu Ala Val Lys Gly Cys
    1650                1655                1660
His Leu Phe Gly Ala Gly Lys Ile Glu Asn Gly Leu Cys Ile Cys Ala
1665                1670                1675                1680
Ala Met Pro Ser Lys Val Val Ile Leu Arg Tyr Asn Glu Asn Leu Ser
                1685                1690                1695
Lys Tyr Cys Ile Arg Lys Glu Ile Glu Thr Ser Glu Pro Cys Ser Cys
            1700                1705                1710
Ile His Phe Thr Asn Tyr Ser Ile Leu Ile Gly Thr Asn Lys Phe Tyr
        1715                1720                1725
Glu Ile Asp Met Lys Gln Tyr Thr Leu Glu Glu Phe Leu Asp Lys Asn
    1730                1735                1740
Asp His Ser Leu Ala Pro Ala Val Phe Ala Ala Ser Ser Asn Ser Phe
1745                1750                1755                1760
Pro Val Ser Ile Val Gln Val Asn Ser Ala Gly Gln Arg Glu Glu Tyr
                1765                1770                1775
Leu Leu Cys Phe His Glu Phe Gly Val Phe Val Asp Ser Tyr Gly Arg
            1780                1785                1790
Arg Ser Arg Thr Asp Asp Leu Lys Trp Ser Arg Leu Pro Leu Ala Phe
        1795                1800                1805
Ala Tyr Arg Glu Pro Tyr Leu Phe Val Thr His Phe Asn Ser Leu Glu
    1810                1815                1820
Val Ile Glu Ile Gln Ala Arg Ser Ser Ala Gly Thr Pro Ala Arg Ala
1825                1830                1835                1840
Tyr Leu Asp Ile Pro Asn Pro Arg Tyr Leu Gly Pro Ala Ile Ser Ser
                1845                1850                1855
Gly Ala Ile Tyr Leu Ala Ser Ser Tyr Gln Asp Lys Leu Arg Val Ile
            1860                1865                1870
Cys Cys Lys Gly Asn Leu Val Lys Glu Ser Gly Thr Glu His His Arg
```

```
        1875                1880                1885
Gly Pro Ser Thr Ser Arg Ser Ser Pro Asn Lys Arg Gly Pro Pro Thr
    1890                1895                1900
Tyr Asn Glu His Ile Thr Lys Arg Val Ala Ser Ser Pro Ala Pro Pro
1905                1910                1915                1920
Glu Gly Pro Ser His Pro Arg Glu Pro Ser Thr Pro His Arg Tyr Arg
                1925                1930                1935
Glu Gly Arg Thr Glu Leu Arg Arg Asp Lys Ser Pro Gly Arg Pro Leu
            1940                1945                1950
Glu Arg Glu Lys Ser Pro Gly Arg Met Leu Ser Thr Arg Arg Glu Arg
        1955                1960                1965
Ser Pro Gly Arg Leu Phe Glu Asp Ser Ser Arg Gly Arg Leu Pro Ala
    1970                1975                1980
Gly Ala Val Arg Thr Pro Leu Ser Gln Val Asn Lys Gly Arg Gly Gln
1985                1990                1995                2000
Ser Ala Ser Gln Val Phe Thr Val Asn Thr Val Thr Tyr Tyr Asp Trp
                2005                2010                2015
Asn Lys Lys Leu Asp Asn Leu Pro Ala Asn Trp Ser Val Leu Arg Ile
            2020                2025                2030
Ile Gln Leu Asn Gly Glu Ile Arg Gln Gln Val Glu Lys Ser Val Leu
        2035                2040                2045
Arg Thr Asp Tyr Cys
    2050
```

<210> 13
<211> 3152
<212> DNA
<213> Homo Sapiens

<400> 13

```
gctcgcgaat gtagcaggcg ccccaagctc ggtcctcaag aagccatggc ggaatccagg 60
ggccgtctgt acctttggat gtgcttggct gctgcgctgg catctttcct gatgggattt 120
atggtgggct ggtttattaa gcctctcaaa gaaacaacca cttctgtgcg ctatcatcaa 180
agtatacggt ggaaactggt atccgaaatg aaagctgaaa acatcaaatc atttcttcgt 240
tcttttacaa agcttcctca tctggcagga acagaacaaa atttcttgct tgccaagaaa 300
atccaaaccc agtggaagaa atttggacta gattcagcca agttggttca ttatgatgtc 360
ctcttatctt accccaatga gacaaatgcc aactatatat cgattgtgga tgaacatgaa 420
actgagattt tcaaaacatc ataccttgaa ccaccaccag atggctatga gaatgttaca 480
aatattgtgc caccatataa tgctttctca gcccaaggca tgccagaggg agatcttgta 540
tatgtgaact atgctcgcac tgaagacttt ttcaaactag aaagagagat gggcatcaac 600
tgtactggga agattgttat tgcaagatat ggaaaaatct tcagaggaaa taaagttaaa 660
aatgccatgt tagcaggagc cataggaatc atcttgtact cagatccagc tgactacttt 720
gctcctgagg tacagccata tcccaaagga tggaatcttc ctggaactgc agcccagaga 780
ggaaatgtgt taaatttgaa tggtgctggt gacccactca ctccaggcta tccagcaaaa 840
gaatacactt tcagacttga tgttgaagaa ggagtgggaa tcccccgaat acctgtacat 900
cccattggat ataatgatgc agaaatatta ttacgctact tgggaggaat tgctccacca 960
gataagagtt ggaagggagc ccttaatgtg agttatagta tcggacctgg ctttacaggg 1020
agtgattctt tcaggaaggt tagaatgcat gtttataaca tcaataaaat tacaaggatt 1080
tacaatgtag ttggaactat cagaggatct gtggaacctg acaggtatgt tattctggga 1140
ggtcaccggg actcctgggt atttggagct attgacccaa ccagtggggt tgctgttttg 1200
caagaaattg cccggagttt tggaaaactg atgagtaaag gctggagacc tagaagaact 1260
atcatttttg ccagctggga tgcagaagaa tttggacttc tgggttccac agaatgggct 1320
gaggagaatg tcaaaatact ccaggagaga agcattgctt atatcaactc ggattcatct 1380
atagaaggca attatactct cagagttgac tgtactcccc ttctttacca attagtgtat 1440
aaactgacaa aagagatccc cagccctgat gatgggtttg agagtaaatc actgtatgaa 1500
agctggttgg aaaaagaccc ttcacctgaa aataaaaatt tgcctagaat caataagctg 1560
ggatctggaa gtgactttga agcttatttt cagagacttg gaattgcttc aggcagagcc 1620
cgttacacta agaataagaa aacagataag tacagcagct acccagtgta ccacacaatt 1680
tatgagacat ttgaattggt agagaaattt tatgacccca catttaaaaa acaactttct 1740
gtggctcaat tacgaggagc actggtatat gagcttgtgg attctaaaat cattcctttt 1800
aatattcaag actatgcaga agctttgaaa aactatgcag caagtatcta taatctatct 1860
```

```
aagaaacatg atcaacaatt aacagaccat ggagtatcat ttgactcctt attttctgct 1920
gtgaaaaact tctcagaggc tgcttcagat tttcataaac gacttataca agttgatctt 1980
aacaatccca ttgcagtgag aatgatgaat gaccaactga tgctcctgga aagagcattc 2040
atcgatcctc ttggtttacc aggaaagctg ttctataggc acatcatatt tgctccaagt 2100
agccacaaca aatatgctgg agaatcattt cctggaatct atgatgctat ctttgatatt 2160
gaaaataaag ccaactctcg tttggcctgg aaagaagtaa agaaacatat ttctattgca 2220
gcttttacaa ttcaagcagc agcaggaact ctgaaagaag tattatagaa ggtctcaagt 2280
ggctagccat taaaggtgtt gctaaaagtc tgaggataaa attcaccttt ctgataactt 2340
atgaagccag ggtgttctaa actcttttca tgtcatgttt tgattatagg ctttggtctt 2400
ttcatctgca aagccttttt ttttttgct ctttaaaagt taataattat attagcaaag 2460
tgttaatcta atgaagtaaa aaactcctgt gtggcagaaa gtaaaagaaa attccctaaa 2520
ttatagcaag gaacatgaat tctcagacat tgtgagtgtg ggaatgtaaa atggtaaaat 2580
cacttttgaa aacagtttgg cagtttccta taaagttaaa catacacttt tactttagga 2640
ctccagaatt ccacttctag ttatttattc aagagaagga aaaacaatga tcacagcaat 2700
acttgtatgc atgttcattg caacttaaaa gcgtaaaaac cccaaatgtc catccacaga 2760
cgaatgtata aactgtggta tccattacac aatagactac ttactactca gcaataaaaa 2820
tgaagtaact ttcaataaat gcaatattat tggcagacat tgttgaagga aaaaagccag 2880
acaaacaact acataaaata tgtttctatt tagatgaagt ggcaaactaa tctgtagtgt 2940
taaaaattag attagtgatt gcctgggcca agtggcaggt tggggaggat ggctgcaaag 3000
aagtatgagg aaactttctc caatagatga gaattttccg tatcttgatc tgagtggcaa 3060
attgtaaact taaaatatat ataaaattta ttgaaagaaa attaagcctc aataaacgtg 3120
attataaaaa aaaaaaaaaa aagggcggcc gc                                3152

<210> 14
<211> 2223
<212> DNA
<213> Homo Sapiens

<220>
<221> CDS
<222> (1)...(2223)

<400> 14
atg gcg gaa tcc agg ggc cgt ctg tac ctt tgg atg tgc ttg gct gct   48
Met Ala Glu Ser Arg Gly Arg Leu Tyr Leu Trp Met Cys Leu Ala Ala
 1               5                  10                  15

gcg ctg gca tct ttc ctg atg gga ttt atg gtg ggc tgg ttt att aag   96
Ala Leu Ala Ser Phe Leu Met Gly Phe Met Val Gly Trp Phe Ile Lys
             20                  25                  30

cct ctc aaa gaa aca acc act tct gtg cgc tat cat caa agt ata cgg   144
Pro Leu Lys Glu Thr Thr Thr Ser Val Arg Tyr His Gln Ser Ile Arg
         35                  40                  45

tgg aaa ctg gta tcc gaa atg aaa gct gaa aac atc aaa tca ttt ctt   192
Trp Lys Leu Val Ser Glu Met Lys Ala Glu Asn Ile Lys Ser Phe Leu
     50                  55                  60

cgt tct ttt aca aag ctt cct cat ctg gca gga aca gaa caa aat ttc   240
Arg Ser Phe Thr Lys Leu Pro His Leu Ala Gly Thr Glu Gln Asn Phe
 65                  70                  75                  80

ttg ctt gcc aag aaa atc caa acc cag tgg aag aaa ttt gga cta gat   288
Leu Leu Ala Lys Lys Ile Gln Thr Gln Trp Lys Lys Phe Gly Leu Asp
                 85                  90                  95

tca gcc aag ttg gtt cat tat gat gtc ctc tta tct tac ccc aat gag   336
Ser Ala Lys Leu Val His Tyr Asp Val Leu Leu Ser Tyr Pro Asn Glu
            100                 105                 110

aca aat gcc aac tat ata tcg att gtg gat gaa cat gaa act gag att   384
```

```
Thr Asn Ala Asn Tyr Ile Ser Ile Val Asp Glu His Glu Thr Glu Ile
        115                 120                 125

ttc aaa aca tca tac ctt gaa cca cca cca gat ggc tat gag aat gtt   432
Phe Lys Thr Ser Tyr Leu Glu Pro Pro Pro Asp Gly Tyr Glu Asn Val
    130                 135                 140

aca aat att gtg cca cca tat aat gct ttc tca gcc caa ggc atg cca   480
Thr Asn Ile Val Pro Pro Tyr Asn Ala Phe Ser Ala Gln Gly Met Pro
145                 150                 155                 160

gag gga gat ctt gta tat gtg aac tat gct cgc act gaa gac ttt ttc   528
Glu Gly Asp Leu Val Tyr Val Asn Tyr Ala Arg Thr Glu Asp Phe Phe
                165                 170                 175

aaa cta gaa aga gag atg ggc atc aac tgt act ggg aag att gtt att   576
Lys Leu Glu Arg Glu Met Gly Ile Asn Cys Thr Gly Lys Ile Val Ile
            180                 185                 190

gca aga tat gga aaa atc ttc aga gga aat aaa gtt aaa aat gcc atg   624
Ala Arg Tyr Gly Lys Ile Phe Arg Gly Asn Lys Val Lys Asn Ala Met
        195                 200                 205

tta gca gga gcc ata gga atc atc ttg tac tca gat cca gct gac tac   672
Leu Ala Gly Ala Ile Gly Ile Ile Leu Tyr Ser Asp Pro Ala Asp Tyr
    210                 215                 220

ttt gct cct gag gta cag cca tat ccc aaa gga tgg aat ctt cct gga   720
Phe Ala Pro Glu Val Gln Pro Tyr Pro Lys Gly Trp Asn Leu Pro Gly
225                 230                 235                 240

act gca gcc cag aga gga aat gtg tta aat ttg aat ggt gct ggt gac   768
Thr Ala Ala Gln Arg Gly Asn Val Leu Asn Leu Asn Gly Ala Gly Asp
                245                 250                 255

cca ctc act cca ggc tat cca gca aaa gaa tac act ttc aga ctt gat   816
Pro Leu Thr Pro Gly Tyr Pro Ala Lys Glu Tyr Thr Phe Arg Leu Asp
            260                 265                 270

gtt gaa gaa gga gtg gga atc ccc cga ata cct gta cat ccc att gga   864
Val Glu Glu Gly Val Gly Ile Pro Arg Ile Pro Val His Pro Ile Gly
        275                 280                 285

tat aat gat gca gaa ata tta tta cgc tac ttg gga gga att gct cca   912
Tyr Asn Asp Ala Glu Ile Leu Leu Arg Tyr Leu Gly Gly Ile Ala Pro
    290                 295                 300

cca gat aag agt tgg aag gga gcc ctt aat gtg agt tat agt atc gga   960
Pro Asp Lys Ser Trp Lys Gly Ala Leu Asn Val Ser Tyr Ser Ile Gly
305                 310                 315                 320

cct ggc ttt aca ggg agt gat tct ttc agg aag gtt aga atg cat gtt   1008
Pro Gly Phe Thr Gly Ser Asp Ser Phe Arg Lys Val Arg Met His Val
                325                 330                 335

tat aac atc aat aaa att aca agg att tac aat gta gtt gga act atc   1056
Tyr Asn Ile Asn Lys Ile Thr Arg Ile Tyr Asn Val Val Gly Thr Ile
            340                 345                 350

aga gga tct gtg gaa cct gac agg tat gtt att ctg gga ggt cac cgg   1104
Arg Gly Ser Val Glu Pro Asp Arg Tyr Val Ile Leu Gly Gly His Arg
        355                 360                 365
```

```
gac tcc tgg gta ttt gga gct att gac cca acc agt ggg gtt gct gtt   1152
Asp Ser Trp Val Phe Gly Ala Ile Asp Pro Thr Ser Gly Val Ala Val
    370                 375                 380

ttg caa gaa att gcc cgg agt ttt gga aaa ctg atg agt aaa ggc tgg   1200
Leu Gln Glu Ile Ala Arg Ser Phe Gly Lys Leu Met Ser Lys Gly Trp
385                 390                 395                 400

aga cct aga aga act atc att ttt gcc agc tgg gat gca gaa gaa ttt   1248
Arg Pro Arg Arg Thr Ile Ile Phe Ala Ser Trp Asp Ala Glu Glu Phe
                405                 410                 415

gga ctt ctg ggt tcc aca gaa tgg gct gag gag aat gtc aaa ata ctc   1296
Gly Leu Leu Gly Ser Thr Glu Trp Ala Glu Glu Asn Val Lys Ile Leu
            420                 425                 430

cag gag aga agc att gct tat atc aac tcg gat tca tct ata gaa ggc   1344
Gln Glu Arg Ser Ile Ala Tyr Ile Asn Ser Asp Ser Ser Ile Glu Gly
        435                 440                 445

aat tat act ctc aga gtt gac tgt act ccc ctt ctt tac caa tta gtg   1392
Asn Tyr Thr Leu Arg Val Asp Cys Thr Pro Leu Leu Tyr Gln Leu Val
    450                 455                 460

tat aaa ctg aca aaa gag atc ccc agc cct gat gat ggg ttt gag agt   1440
Tyr Lys Leu Thr Lys Glu Ile Pro Ser Pro Asp Asp Gly Phe Glu Ser
465                 470                 475                 480

aaa tca ctg tat gaa agc tgg ttg gaa aaa gac cct tca cct gaa aat   1488
Lys Ser Leu Tyr Glu Ser Trp Leu Glu Lys Asp Pro Ser Pro Glu Asn
                485                 490                 495

aaa aat ttg cct aga atc aat aag ctg gga tct gga agt gac ttt gaa   1536
Lys Asn Leu Pro Arg Ile Asn Lys Leu Gly Ser Gly Ser Asp Phe Glu
            500                 505                 510

gct tat ttt cag aga ctt gga att gct tca ggc aga gcc cgt tac act   1584
Ala Tyr Phe Gln Arg Leu Gly Ile Ala Ser Gly Arg Ala Arg Tyr Thr
        515                 520                 525

aag aat aag aaa aca gat aag tac agc agc tac cca gtg tac cac aca   1632
Lys Asn Lys Lys Thr Asp Lys Tyr Ser Ser Tyr Pro Val Tyr His Thr
    530                 535                 540

att tat gag aca ttt gaa ttg gta gag aaa ttt tat gac ccc aca ttt   1680
Ile Tyr Glu Thr Phe Glu Leu Val Glu Lys Phe Tyr Asp Pro Thr Phe
545                 550                 555                 560

aaa aaa caa ctt tct gtg gct caa tta cga gga gca ctg gta tat gag   1728
Lys Lys Gln Leu Ser Val Ala Gln Leu Arg Gly Ala Leu Val Tyr Glu
                565                 570                 575

ctt gtg gat tct aaa atc att cct ttt aat att caa gac tat gca gaa   1776
Leu Val Asp Ser Lys Ile Ile Pro Phe Asn Ile Gln Asp Tyr Ala Glu
            580                 585                 590

gct ttg aaa aac tat gca gca agt atc tat aat cta tct aag aaa cat   1824
Ala Leu Lys Asn Tyr Ala Ala Ser Ile Tyr Asn Leu Ser Lys Lys His
        595                 600                 605

gat caa caa tta aca gac cat gga gta tca ttt gac tcc tta ttt tct   1872
```

```
Asp Gln Gln Leu Thr Asp His Gly Val Ser Phe Asp Ser Leu Phe Ser
    610                 615                 620

gct gtg aaa aac ttc tca gag gct gct tca gat ttt cat aaa cga ctt   1920
Ala Val Lys Asn Phe Ser Glu Ala Ala Ser Asp Phe His Lys Arg Leu
625                 630                 635                 640

ata caa gtt gat ctt aac aat ccc att gca gtg aga atg atg aat gac   1968
Ile Gln Val Asp Leu Asn Asn Pro Ile Ala Val Arg Met Met Asn Asp
                645                 650                 655

caa ctg atg ctc ctg gaa aga gca ttc atc gat cct ctt ggt tta cca   2016
Gln Leu Met Leu Leu Glu Arg Ala Phe Ile Asp Pro Leu Gly Leu Pro
            660                 665                 670

gga aag ctg ttc tat agg cac atc ata ttt gct cca agt agc cac aac   2064
Gly Lys Leu Phe Tyr Arg His Ile Ile Phe Ala Pro Ser Ser His Asn
        675                 680                 685

aaa tat gct gga gaa tca ttt cct gga atc tat gat gct atc ttt gat   2112
Lys Tyr Ala Gly Glu Ser Phe Pro Gly Ile Tyr Asp Ala Ile Phe Asp
    690                 695                 700

att gaa aat aaa gcc aac tct cgt ttg gcc tgg aaa gaa gta aag aaa   2160
Ile Glu Asn Lys Ala Asn Ser Arg Leu Ala Trp Lys Glu Val Lys Lys
705                 710                 715                 720

cat att tct att gca gct ttt aca att caa gca gca gca gga act ctg   2208
His Ile Ser Ile Ala Ala Phe Thr Ile Gln Ala Ala Ala Gly Thr Leu
                725                 730                 735

aaa gaa gta tta tag                                             2223
Lys Glu Val Leu  *
            740
```

<210> 15
<211> 740
<212> PRT
<213> Homo Sapiens

<400> 15

```
Met Ala Glu Ser Arg Gly Arg Leu Tyr Leu Trp Met Cys Leu Ala Ala
 1               5                  10                  15
Ala Leu Ala Ser Phe Leu Met Gly Phe Met Val Gly Trp Phe Ile Lys
            20                  25                  30
Pro Leu Lys Glu Thr Thr Thr Ser Val Arg Tyr His Gln Ser Ile Arg
        35                  40                  45
Trp Lys Leu Val Ser Glu Met Lys Ala Glu Asn Ile Lys Ser Phe Leu
    50                  55                  60
Arg Ser Phe Thr Lys Leu Pro His Leu Ala Gly Thr Glu Gln Asn Phe
65                  70                  75                  80
Leu Leu Ala Lys Lys Ile Gln Thr Gln Trp Lys Lys Phe Gly Leu Asp
                85                  90                  95
Ser Ala Lys Leu Val His Tyr Asp Val Leu Leu Ser Tyr Pro Asn Glu
            100                 105                 110
Thr Asn Ala Asn Tyr Ile Ser Ile Val Asp Glu His Glu Thr Glu Ile
        115                 120                 125
Phe Lys Thr Ser Tyr Leu Glu Pro Pro Pro Asp Gly Tyr Glu Asn Val
    130                 135                 140
Thr Asn Ile Val Pro Pro Tyr Asn Ala Phe Ser Ala Gln Gly Met Pro
145                 150                 155                 160
```

```
Glu Gly Asp Leu Val Tyr Val Asn Tyr Ala Arg Thr Glu Asp Phe Phe
                165                 170                 175
Lys Leu Glu Arg Glu Met Gly Ile Asn Cys Thr Gly Lys Ile Val Ile
            180                 185                 190
Ala Arg Tyr Gly Lys Ile Phe Arg Gly Asn Lys Val Lys Asn Ala Met
        195                 200                 205
Leu Ala Gly Ala Ile Gly Ile Ile Leu Tyr Ser Asp Pro Ala Asp Tyr
    210                 215                 220
Phe Ala Pro Glu Val Gln Pro Tyr Pro Lys Gly Trp Asn Leu Pro Gly
225                 230                 235                 240
Thr Ala Ala Gln Arg Gly Asn Val Leu Asn Leu Asn Gly Ala Gly Asp
                245                 250                 255
Pro Leu Thr Pro Gly Tyr Pro Ala Lys Glu Tyr Thr Phe Arg Leu Asp
            260                 265                 270
Val Glu Glu Gly Val Gly Ile Pro Arg Ile Pro Val His Pro Ile Gly
        275                 280                 285
Tyr Asn Asp Ala Glu Ile Leu Leu Arg Tyr Leu Gly Gly Ile Ala Pro
    290                 295                 300
Pro Asp Lys Ser Trp Lys Gly Ala Leu Asn Val Ser Tyr Ser Ile Gly
305                 310                 315                 320
Pro Gly Phe Thr Gly Ser Asp Ser Phe Arg Lys Val Arg Met His Val
                325                 330                 335
Tyr Asn Ile Asn Lys Ile Thr Arg Ile Tyr Asn Val Val Gly Thr Ile
            340                 345                 350
Arg Gly Ser Val Glu Pro Asp Arg Tyr Val Ile Leu Gly Gly His Arg
        355                 360                 365
Asp Ser Trp Val Phe Gly Ala Ile Asp Pro Thr Ser Gly Val Ala Val
    370                 375                 380
Leu Gln Glu Ile Ala Arg Ser Phe Gly Lys Leu Met Ser Lys Gly Trp
385                 390                 395                 400
Arg Pro Arg Arg Thr Ile Ile Phe Ala Ser Trp Asp Ala Glu Glu Phe
                405                 410                 415
Gly Leu Leu Gly Ser Thr Glu Trp Ala Glu Glu Asn Val Lys Ile Leu
            420                 425                 430
Gln Glu Arg Ser Ile Ala Tyr Ile Asn Ser Asp Ser Ser Ile Glu Gly
        435                 440                 445
Asn Tyr Thr Leu Arg Val Asp Cys Thr Pro Leu Leu Tyr Gln Leu Val
    450                 455                 460
Tyr Lys Leu Thr Lys Glu Ile Pro Ser Pro Asp Asp Gly Phe Glu Ser
465                 470                 475                 480
Lys Ser Leu Tyr Glu Ser Trp Leu Glu Lys Asp Pro Ser Pro Glu Asn
                485                 490                 495
Lys Asn Leu Pro Arg Ile Asn Lys Leu Gly Ser Gly Ser Asp Phe Glu
            500                 505                 510
Ala Tyr Phe Gln Arg Leu Gly Ile Ala Ser Gly Arg Ala Arg Tyr Thr
        515                 520                 525
Lys Asn Lys Lys Thr Asp Lys Tyr Ser Ser Tyr Pro Val Tyr His Thr
    530                 535                 540
Ile Tyr Glu Thr Phe Glu Leu Val Glu Lys Phe Tyr Asp Pro Thr Phe
545                 550                 555                 560
Lys Lys Gln Leu Ser Val Ala Gln Leu Arg Gly Ala Leu Val Tyr Glu
                565                 570                 575
Leu Val Asp Ser Lys Ile Ile Pro Phe Asn Ile Gln Asp Tyr Ala Glu
            580                 585                 590
Ala Leu Lys Asn Tyr Ala Ala Ser Ile Tyr Asn Leu Ser Lys Lys His
        595                 600                 605
Asp Gln Gln Leu Thr Asp His Gly Val Ser Phe Asp Ser Leu Phe Ser
    610                 615                 620
Ala Val Lys Asn Phe Ser Glu Ala Ala Ser Asp Phe His Lys Arg Leu
625                 630                 635                 640
Ile Gln Val Asp Leu Asn Asn Pro Ile Ala Val Arg Met Met Asn Asp
                645                 650                 655
```

```
Gln Leu Met Leu Leu Glu Arg Ala Phe Ile Asp Pro Leu Gly Leu Pro
            660                 665                 670
Gly Lys Leu Phe Tyr Arg His Ile Ile Phe Ala Pro Ser Ser His Asn
        675                 680                 685
Lys Tyr Ala Gly Glu Ser Phe Pro Gly Ile Tyr Asp Ala Ile Phe Asp
    690                 695                 700
Ile Glu Asn Lys Ala Asn Ser Arg Leu Ala Trp Lys Glu Val Lys Lys
705                 710                 715                 720
His Ile Ser Ile Ala Ala Phe Thr Ile Gln Ala Ala Ala Gly Thr Leu
                725                 730                 735
Lys Glu Val Leu
            740
```

<210> 16
<211> 3693
<212> DNA
<213> Homo Sapiens

<400> 16

```
aactgtagtg cattttgcct tctttccttc ttagagtcac ctctactttc ctgctaccgc 60
tgcctgtgag ctgaaggggc tgaaccatac actccttttt ctacaaccag cttgcatttt 120
ttctgcccac aatgagcggg gtaagatttt tatttttggc aaggggtata atttgggttc 180
actgtggcta cttgaacact acactgcagc taactctatc tttgtttcct ttccaggaat 240
caatgaattt cagcgatgtt ttcgactcca gtgaagatta ttttgtgtca gtcaatactt 300
catattactc agttgattct gagatgttac tgtgctcctt gcaggaggtc aggcagttct 360
ccaggctatt tgtaccgatt gcctactcct tgatctgtgt ctttggcctc ctggggaata 420
ttctggtggt gatcaccttt gctttttata agaaggccag gtctatgaca gacgtctatc 480
tcttgaacat ggccattgca gacatcctct ttgttcttac tctcccattc tgggcagtga 540
gtcatgccac tggtgcgtgg gttttcagca atgccacgtg caagttgcta aaaggcatct 600
atgccatcaa ctttaactgc gggatgctgc tcctgacttg cattagcatg gaccggtaca 660
tcgccattgt acaggcgact aagtcattcc ggctccgatc cagaacacta ccgcgcagca 720
aaatcatctg ccttgttgtg tgggggctgt cagtcatcat ctccagctca acttttgtct 780
tcaaccaaaa atacaacacc caaggcagcg atgtctgtga acccaagtac cagactgtct 840
cggagcccat caggtggaag ctgctgatgt tggggcttga gctactcttt ggtttcttta 900
tccctttgat gttcatgata ttttgttaca cgttcattgt caaaaccttg gtgcaagctc 960
agaattctaa aaggcacaaa gccatccgtg taatcatagc tgtggtgctt gtgtttctgg 1020
cttgtcagat tcctcataac atggtcctgc ttgtgacggc tgcaaatttg ggtaaaatga 1080
accgatcctg ccagagcgaa aagctaattg gctatacgaa aactgtcaca gaagtcctgg 1140
ctttcctgca ctgctgcctg aaccctgtgc tctacgcttt tattgggcag aagttcagaa 1200
actactttct gaagatcttg aaggacctgt ggtgtgtgag aaggaagtac aagtcctcag 1260
gcttctcctg tgccgggagg tactcagaaa acatttctcg gcagaccagt gagaccgcag 1320
ataacgacaa tgcgtcgtcc ttcactatgt gatagaaagc tgagtctccc taaggcatgt 1380
gtgaaacata ctcatagatg ttatgcaaaa aaaagtctat ggccaggtat gcatggaaaa 1440
tgtgggaatt aagcaaaatc aagcaagcct ctctcctgcg ggacttaacg tgctcatggg 1500
ctgtgtgatc tcttcagggt ggggtggtct ctgataggta gcattttcca gcactttgca 1560
aggaatgttt tgtagctcta gggtatatat ccgcctggca tttcacaaaa cagcctttgg 1620
gaaatgctga attaaagtga attgttgaca aatgtaaaca ttttcagaaa tattcatgaa 1680
gcggtcacag atcacagtgt cttttggtta cagcacaaaa tgatggcagt ggtttgaaaa 1740
actaaaacag aaaaaaaaat ggaagccaac acatcactca ttttaggcaa atgtttaaac 1800
atttttatct atcagaatgt ttattgttgc tggttataag cagcaggatt ggccggctag 1860
tgtttcctct catttccctt tgatacagtc aacaagcctg accctgtaaa atggaggtgg 1920
aaagacaagc tcaagtgttc acaacctgga agtgcttcgg aaagaagggg acaatggcag 1980
aacaggtgtt ggtgacaatt gtcaccaatt ggataaagca gctcaggttg tagtgggcca 2040
ttaggaaact gtcggtttgc tttgatttcc ctgggagctg ttctctgtcg tgagtgtctc 2100
ttgtctaaac gtccattaag ctgagagtgc tatgaagaca ggatctagaa taatcttgct 2160
cacagctgtg ctctgagtgc ctagcggagt tccagcaaac aaaatggact caagagagat 2220
ttgattaatg aatcgtaatg aagttggggt ttattgtaca gtttaaaatg ttagatgttt 2280
ttaatttttt aaataaatgg aatacttttt tttttttta aagaaagcaa ctttactgag 2340
acaatgtaga aagaagtttt gttccgtttc tttaatgtgg ttgaagagca atgtgtggct 2400
gaagactttt gttatgagga gctgcagatt agctagggga cagctggaat tatgctggct 2460
tctgataatt attttaaagg ggtctgaaat ttgtgatgga atcagatttt aacagctctc 2520
```

```
ttcaatgaca tagaaagttc atggaactca tgtttttaaa gggctatgta aatatatgaa 2580
cattagaaaa atagcaactt gtgttacaaa aatacaaaca catgttagga aggtactgtc 2640
atgggctagg catggtggct cacacctgta atcccagcat tttgggaagc taagatgggt 2700
ggatcacttg aggtcaggag tttgagacca gcctggccaa catggcgaaa ccccctctca 2760
ctaaaaatac aaaaatttgc caggcgtggt ggcgggtgcc tgtaatccca gctacttggg 2820
aggctgaggc aagagaatcg cttgaaccca ggaggcagag gttgcagtga gccgagatcg 2880
tgccattgca ctccagcctg ggtgacaaag cgagactcca tctcaaaaaa aaaaaaaaaa 2940
aaaaaaggaa agaactgtca tgtaaacata ccgacatgtt taaacctgac aatggtgtta 3000
tttgaaactt tatattgttc ttgtaagctt taactatatc tctctttaaa atgcaaaata 3060
atgtcttaag attcaaagtc tgtattttta aagcatggct ttggctttgc aaaataaaaa 3120
atgtgttttg tacatgaagt aggaatcgta tttcagcttc aaggttcaga ttgaggggcc 3180
cactgtttgg agaggatggt attcaggctt tctcatgtcc ttcaaatctg ttagcgtttg 3240
actctagaaa tcaaagcaaa ggagtggtta cccagacact tcttttggtg tgatcaatgc 3300
gctgatgtga tctatgaaga tgattcatgc ttgaaaacta gcacagaaac atcttgctta 3360
tttgccaaag ctgggagatg agcttctctg cataatttaa atgttcagat aaatgaagct 3420
gacttattta agcaataacc ttttaaacat tttagctaag atgtataaaa atgtttccaa 3480
aatataccac atactttatt tcttcttaaa tgtagtacat taggttacat cattttttctt 3540
gctgtcttgg gcatcaaaac aggtgccatg gtaacctgac actctcagga gacattaaga 3600
tagaaggggc tgttcttcag tggttcccat tgattctccc catatctttt tgctctcagg 3660
ctctggccgt ctcttcctga gccttaactg tgt                              3693

<210> 17
<211> 1221
<212> DNA
<213> Homo Sapiens

<220>
<221> CDS
<222> (1)...(1221)

<400> 17
atg agc ggg gta aga ttt tta ttt ttg gca agg ggt ata att tgg gtt   48
Met Ser Gly Val Arg Phe Leu Phe Leu Ala Arg Gly Ile Ile Trp Val
 1               5                  10                  15

cac tgt ggc tac ttg aac act aca ctg cag cta act cta tct ttg ttt   96
His Cys Gly Tyr Leu Asn Thr Thr Leu Gln Leu Thr Leu Ser Leu Phe
             20                  25                  30

cct ttc cag gaa tca atg aat ttc agc gat gtt ttc gac tcc agt gaa   144
Pro Phe Gln Glu Ser Met Asn Phe Ser Asp Val Phe Asp Ser Ser Glu
         35                  40                  45

gat tat ttt gtg tca gtc aat act tca tat tac tca gtt gat tct gag   192
Asp Tyr Phe Val Ser Val Asn Thr Ser Tyr Tyr Ser Val Asp Ser Glu
     50                  55                  60

atg tta ctg tgc tcc ttg cag gag gtc agg cag ttc tcc agg cta ttt   240
Met Leu Leu Cys Ser Leu Gln Glu Val Arg Gln Phe Ser Arg Leu Phe
 65                  70                  75                  80

gta ccg att gcc tac tcc ttg atc tgt gtc ttt ggc ctc ctg ggg aat   288
Val Pro Ile Ala Tyr Ser Leu Ile Cys Val Phe Gly Leu Leu Gly Asn
                 85                  90                  95

att ctg gtg gtg atc acc ttt gct ttt tat aag aag gcc agg tct atg   336
Ile Leu Val Val Ile Thr Phe Ala Phe Tyr Lys Lys Ala Arg Ser Met
            100                 105                 110

aca gac gtc tat ctc ttg aac atg gcc att gca gac atc ctc ttt gtt   384
Thr Asp Val Tyr Leu Leu Asn Met Ala Ile Ala Asp Ile Leu Phe Val
        115                 120                 125
```

```
ctt act ctc cca ttc tgg gca gtg agt cat gcc act ggt gcg tgg gtt   432
Leu Thr Leu Pro Phe Trp Ala Val Ser His Ala Thr Gly Ala Trp Val
    130                 135                 140

ttc agc aat gcc acg tgc aag ttg cta aaa ggc atc tat gcc atc aac   480
Phe Ser Asn Ala Thr Cys Lys Leu Leu Lys Gly Ile Tyr Ala Ile Asn
145                 150                 155                 160

ttt aac tgc ggg atg ctg ctc ctg act tgc att agc atg gac cgg tac   528
Phe Asn Cys Gly Met Leu Leu Leu Thr Cys Ile Ser Met Asp Arg Tyr
                165                 170                 175

atc gcc att gta cag gcg act aag tca ttc cgg ctc cga tcc aga aca   576
Ile Ala Ile Val Gln Ala Thr Lys Ser Phe Arg Leu Arg Ser Arg Thr
            180                 185                 190

cta ccg cgc agc aaa atc atc tgc ctt gtt gtg tgg ggg ctg tca gtc   624
Leu Pro Arg Ser Lys Ile Ile Cys Leu Val Val Trp Gly Leu Ser Val
        195                 200                 205

atc atc tcc agc tca act ttt gtc ttc aac caa aaa tac aac acc caa   672
Ile Ile Ser Ser Ser Thr Phe Val Phe Asn Gln Lys Tyr Asn Thr Gln
    210                 215                 220

ggc agc gat gtc tgt gaa ccc aag tac cag act gtc tcg gag ccc atc   720
Gly Ser Asp Val Cys Glu Pro Lys Tyr Gln Thr Val Ser Glu Pro Ile
225                 230                 235                 240

agg tgg aag ctg ctg atg ttg ggg ctt gag cta ctc ttt ggt ttc ttt   768
Arg Trp Lys Leu Leu Met Leu Gly Leu Glu Leu Leu Phe Gly Phe Phe
                245                 250                 255

atc cct ttg atg ttc atg ata ttt tgt tac acg ttc att gtc aaa acc   816
Ile Pro Leu Met Phe Met Ile Phe Cys Tyr Thr Phe Ile Val Lys Thr
            260                 265                 270

ttg gtg caa gct cag aat tct aaa agg cac aaa gcc atc cgt gta atc   864
Leu Val Gln Ala Gln Asn Ser Lys Arg His Lys Ala Ile Arg Val Ile
        275                 280                 285

ata gct gtg gtg ctt gtg ttt ctg gct tgt cag att cct cat aac atg   912
Ile Ala Val Val Leu Val Phe Leu Ala Cys Gln Ile Pro His Asn Met
    290                 295                 300

gtc ctg ctt gtg acg gct gca aat ttg ggt aaa atg aac cga tcc tgc   960
Val Leu Leu Val Thr Ala Ala Asn Leu Gly Lys Met Asn Arg Ser Cys
305                 310                 315                 320

cag agc gaa aag cta att ggc tat acg aaa act gtc aca gaa gtc ctg   1008
Gln Ser Glu Lys Leu Ile Gly Tyr Thr Lys Thr Val Thr Glu Val Leu
                325                 330                 335

gct ttc ctg cac tgc tgc ctg aac cct gtg ctc tac gct ttt att ggg   1056
Ala Phe Leu His Cys Cys Leu Asn Pro Val Leu Tyr Ala Phe Ile Gly
            340                 345                 350

cag aag ttc aga aac tac ttt ctg aag atc ttg aag gac ctg tgg tgt   1104
Gln Lys Phe Arg Asn Tyr Phe Leu Lys Ile Leu Lys Asp Leu Trp Cys
        355                 360                 365

gtg aga agg aag tac aag tcc tca ggc ttc tcc tgt gcc ggg agg tac   1152
```

```
Val Arg Arg Lys Tyr Lys Ser Ser Gly Phe Ser Cys Ala Gly Arg Tyr
    370                 375                 380

tca gaa aac att tct cgg cag acc agt gag acc gca gat aac gac aat   1200
Ser Glu Asn Ile Ser Arg Gln Thr Ser Glu Thr Ala Asp Asn Asp Asn
385                 390                 395                 400

gcg tcg tcc ttc act atg tga                                     1221
Ala Ser Ser Phe Thr Met  *
                405
```

<210> 18
<211> 406
<212> PRT
<213> Homo Sapiens

<400> 18

```
Met Ser Gly Val Arg Phe Leu Phe Leu Ala Arg Gly Ile Ile Trp Val
 1               5                  10                  15
His Cys Gly Tyr Leu Asn Thr Thr Leu Gln Leu Thr Leu Ser Leu Phe
            20                  25                  30
Pro Phe Gln Glu Ser Met Asn Phe Ser Asp Val Phe Asp Ser Ser Glu
        35                  40                  45
Asp Tyr Phe Val Ser Val Asn Thr Ser Tyr Tyr Ser Val Asp Ser Glu
    50                  55                  60
Met Leu Leu Cys Ser Leu Gln Glu Val Arg Gln Phe Ser Arg Leu Phe
65                  70                  75                  80
Val Pro Ile Ala Tyr Ser Leu Ile Cys Val Phe Gly Leu Leu Gly Asn
                85                  90                  95
Ile Leu Val Val Ile Thr Phe Ala Phe Tyr Lys Lys Ala Arg Ser Met
            100                 105                 110
Thr Asp Val Tyr Leu Leu Asn Met Ala Ile Ala Asp Ile Leu Phe Val
        115                 120                 125
Leu Thr Leu Pro Phe Trp Ala Val Ser His Ala Thr Gly Ala Trp Val
    130                 135                 140
Phe Ser Asn Ala Thr Cys Lys Leu Leu Lys Gly Ile Tyr Ala Ile Asn
145                 150                 155                 160
Phe Asn Cys Gly Met Leu Leu Leu Thr Cys Ile Ser Met Asp Arg Tyr
                165                 170                 175
Ile Ala Ile Val Gln Ala Thr Lys Ser Phe Arg Leu Arg Ser Arg Thr
            180                 185                 190
Leu Pro Arg Ser Lys Ile Ile Cys Leu Val Val Trp Gly Leu Ser Val
        195                 200                 205
Ile Ile Ser Ser Ser Thr Phe Val Phe Asn Gln Lys Tyr Asn Thr Gln
    210                 215                 220
Gly Ser Asp Val Cys Glu Pro Lys Tyr Gln Thr Val Ser Glu Pro Ile
225                 230                 235                 240
Arg Trp Lys Leu Leu Met Leu Gly Leu Glu Leu Leu Phe Gly Phe Phe
                245                 250                 255
Ile Pro Leu Met Phe Met Ile Phe Cys Tyr Thr Phe Ile Val Lys Thr
            260                 265                 270
Leu Val Gln Ala Gln Asn Ser Lys Arg His Lys Ala Ile Arg Val Ile
        275                 280                 285
Ile Ala Val Val Leu Val Phe Leu Ala Cys Gln Ile Pro His Asn Met
    290                 295                 300
Val Leu Leu Val Thr Ala Ala Asn Leu Gly Lys Met Asn Arg Ser Cys
305                 310                 315                 320
Gln Ser Glu Lys Leu Ile Gly Tyr Thr Lys Thr Val Thr Glu Val Leu
                325                 330                 335
Ala Phe Leu His Cys Cys Leu Asn Pro Val Leu Tyr Ala Phe Ile Gly
            340                 345                 350
```

```
Gln Lys Phe Arg Asn Tyr Phe Leu Lys Ile Leu Lys Asp Leu Trp Cys
        355                 360                 365
Val Arg Arg Lys Tyr Lys Ser Ser Gly Phe Ser Cys Ala Gly Arg Tyr
    370                 375                 380
Ser Glu Asn Ile Ser Arg Gln Thr Ser Glu Thr Ala Asp Asn Asp Asn
385                 390                 395                 400
Ala Ser Ser Phe Thr Met
                405
```

<210> 19
<211> 3164
<212> DNA
<213> Homo Sapiens

<400> 19

```
atgccggtcc gcaggggcca cgtcgctccc caaaacactt acctggacac catcatccgc 60
aagttcgagg gccaaagtcg gaagttcctg attgccaatg ctcagatgga gaactgcgcc 120
atcatttact gcaacgacgg cttctgcgaa ctcttcggct actcccgagt ggaggtgatg 180
cagcaaccct gcacctgcga cttcctcaca ggccccaaca caccaagcag cgccgtgtcc 240
cgcctagcgc aggccctgct gggggctgag gagtgcaagg tggacatcct ctactaccgc 300
aaggatgcct ccagcttccg ctgcctggta gatgtggtgc ccgtgaagaa cgaggacggg 360
gctgtcatca tgttcattct caacttcgag gacctggccc agctcctggc caagtgcagc 420
agccgcagct tgtcccagcg cctgttgtcc cagagcttcc tgggctccga gggctctcat 480
ggcaggccag gcggaccagg gccaggcaca ggcaggggca agtacaggac catcagccag 540
atcccacagt tcacgctcaa cttcgtggag ttcaacttgg agaagcaccg ctccagctcc 600
accacggaga ttgagatcat cgcgccccat aaggtggtgg agcggacaca gaacgtcact 660
gagaaggtca cccaggtcct gtccctgggc gcggatgtgc tgccggagta caagctgcag 720
gcgccgcgca tccaccgctg gaccatcctg cactacagcc ccttcaaggc cgtgtgggac 780
tggctcatcc tgctgctggt catctacacg gctgtcttca cgccctactc agccgccttc 840
ctgctcagcg accaggacga atcacggcgt ggggcctgca gctatacctg cagtcccctc 900
actgtggtgg atctcatcgt ggacatcatg ttcgtcgtgg acatcgtcat caacttccgc 960
accacctatg tcaacaccaa tgatgaggtg gtcagccacc cccgccgcat cgccgtccac 1020
tacttcaagg gctggttcct cattgacatg gtggccgcca tccctttcga cctcctgatc 1080
ttccgcactg gctccgatga gaccacaacc ctgattgggc tattgaagac agcgcggctg 1140
ctgcggctgg tgcgcgtagc acggaagctg gaccgctact ctgagtatgg ggcggctgtg 1200
ctcttcttgc tcatgtgcac cttcgcgctc atagcgcact ggctggcctg catctggtac 1260
gccatcggca atgtggagcg gccctaccta gaacacaaga tcggctggct ggacagcctg 1320
ggtgtgcagc ttggcaagcg ctacaacggc agcgacccag cctcgggccc ctcggtgcag 1380
gacaagtatg tcacagccct ctacttcacc ttcagcagcc tcaccagcgt gggcttcggc 1440
aatgtctcgc ccaacaccaa ctccgagaag gtcttctcca tctgcgtcat gctcatcggc 1500
tccctgatgt acgccagcat cttcgggaac gtgtccgcga tcatccagcg cctgtactcg 1560
ggcaccgcgc gctaccacac gcagatgctg cgtgtcaagg agttcatccg cttccaccag 1620
atccccaacc cactgcgcca gcgcctggag gagtatttcc agcacgcctg gtcctacacc 1680
aatggcattg acatgaacgc ggtgctgaag ggcttccccg agtgcctgca ggctgacatc 1740
tgcctgcacc tgcaccgcgc actgctgcag cactgcccag ctttcagcgg cgccggcaag 1800
ggctgcctgc gcgcgctagc cgtcaagttc aagaccaccc acgcgccgcc tggggacacg 1860
ctggtgcacc tcggcgacgt gctctccacc ctctacttca tctcccgagg ctccatcgag 1920
atcctgcgcg acgacgtggt cgtggccatc ctaggaaaga atgacatctt tggggaaccc 1980
gtcagcctcc atgcccagcc aggcaagtcc agtgcagacg tgcgggctct gacctactgc 2040
gacctgcaca agatccagcg ggcagatctg ctggaggtgc tggacatgta cccggccttt 2100
gcggagagct tctggagtaa gctggaggtc accttcaacc tgcgggacgc agccgggggt 2160
ctccactcat ccccccgaca ggctcctggc agccaagacc accaaggttt ctttctcagt 2220
gacaaccagt cagatgcagc ccctcccctg agcatctcag atgcatctgg cctctggcct 2280
gagctactgc aggaaatgcc cccaaggcac agcccccaaa gccctcagga agacccagat 2340
tgctggcctc tgaagctggg ctccaggcta gagcagctcc aggcccagat gaacaggctg 2400
gagtcccgcg tgtcctcaga cctcagccgc atcttgcagc tcctccagaa gcccatgccc 2460
cagggccacg ccagctacat tctggaagcc cctgcctcca atgacctggc cttggttcct 2520
atagcctcgg agacgacgag tccagggccc aggctgcccc agggctttct gcctcctgca 2580
cagaccccaa gctatggaga cttggatgac tgtagtccaa agcacaggaa ctcctccccc 2640
aggatgcctc acctggctgt ggcaacggac aaaactctgg caccatcctc agaacaggaa 2700
cagcctgagg ggctctggcc acccctagcc tcacctctac atcccctgga agtacaagga 2760
```

```
ctcatctgtg gtccctgctt ctcctccctc cctgaacacc ttggctctgt tcccaagcag 2820
ctggacttcc agagacatgg ctcagatcct ggatttgcag ggagttgggg ccactgaact 2880
ccaagataaa gacaccatga ggggactgaa ggtgggcaag gggatttcct ttagctgggc 2940
atggtggcgg gcgcctgtaa tcccagctac tcaggaggct gaagcaagag aatcacttga 3000
accctaaagg cagaggttgc agtgagccga gatagtgcca ctgcactaca gcccgggcga 3060
cagagtgaga ctccatctca aaaataaaat aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3120
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa                  3164

<210> 20
<211> 2877
<212> DNA
<213> Homo Sapiens

<220>
<221> CDS
<222> (1)...(2877)

<400> 20
atg ccg gtc cgc agg ggc cac gtc gct ccc caa aac act tac ctg gac   48
Met Pro Val Arg Arg Gly His Val Ala Pro Gln Asn Thr Tyr Leu Asp
 1               5                  10                  15

acc atc atc cgc aag ttc gag ggc caa agt cgg aag ttc ctg att gcc   96
Thr Ile Ile Arg Lys Phe Glu Gly Gln Ser Arg Lys Phe Leu Ile Ala
            20                  25                  30

aat gct cag atg gag aac tgc gcc atc att tac tgc aac gac ggc ttc   144
Asn Ala Gln Met Glu Asn Cys Ala Ile Ile Tyr Cys Asn Asp Gly Phe
        35                  40                  45

tgc gaa ctc ttc ggc tac tcc cga gtg gag gtg atg cag caa ccc tgc   192
Cys Glu Leu Phe Gly Tyr Ser Arg Val Glu Val Met Gln Gln Pro Cys
    50                  55                  60

acc tgc gac ttc ctc aca ggc ccc aac aca cca agc agc gcc gtg tcc   240
Thr Cys Asp Phe Leu Thr Gly Pro Asn Thr Pro Ser Ser Ala Val Ser
65                  70                  75                  80

cgc cta gcg cag gcc ctg ctg ggg gct gag gag tgc aag gtg gac atc   288
Arg Leu Ala Gln Ala Leu Leu Gly Ala Glu Glu Cys Lys Val Asp Ile
                85                  90                  95

ctc tac tac cgc aag gat gcc tcc agc ttc cgc tgc ctg gta gat gtg   336
Leu Tyr Tyr Arg Lys Asp Ala Ser Ser Phe Arg Cys Leu Val Asp Val
            100                 105                 110

gtg ccc gtg aag aac gag gac ggg gct gtc atc atg ttc att ctc aac   384
Val Pro Val Lys Asn Glu Asp Gly Ala Val Ile Met Phe Ile Leu Asn
        115                 120                 125

ttc gag gac ctg gcc cag ctc ctg gcc aag tgc agc agc cgc agc ttg   432
Phe Glu Asp Leu Ala Gln Leu Leu Ala Lys Cys Ser Ser Arg Ser Leu
    130                 135                 140

tcc cag cgc ctg ttg tcc cag agc ttc ctg ggc tcc gag ggc tct cat   480
Ser Gln Arg Leu Leu Ser Gln Ser Phe Leu Gly Ser Glu Gly Ser His
145                 150                 155                 160

ggc agg cca ggc gga cca ggg cca ggc aca ggc agg ggc aag tac agg   528
Gly Arg Pro Gly Gly Pro Gly Pro Gly Thr Gly Arg Gly Lys Tyr Arg
                165                 170                 175
```

```
acc atc agc cag atc cca cag ttc acg ctc aac ttc gtg gag ttc aac   576
Thr Ile Ser Gln Ile Pro Gln Phe Thr Leu Asn Phe Val Glu Phe Asn
            180                 185                 190

ttg gag aag cac cgc tcc agc tcc acc acg gag att gag atc atc gcg   624
Leu Glu Lys His Arg Ser Ser Ser Thr Thr Glu Ile Glu Ile Ile Ala
        195                 200                 205


ccc cat aag gtg gtg gag cgg aca cag aac gtc act gag aag gtc acc   672
Pro His Lys Val Val Glu Arg Thr Gln Asn Val Thr Glu Lys Val Thr
    210                 215                 220

cag gtc ctg tcc ctg ggc gcg gat gtg ctg ccg gag tac aag ctg cag   720
Gln Val Leu Ser Leu Gly Ala Asp Val Leu Pro Glu Tyr Lys Leu Gln
225                 230                 235                 240

gcg ccg cgc atc cac cgc tgg acc atc ctg cac tac agc ccc ttc aag   768
Ala Pro Arg Ile His Arg Trp Thr Ile Leu His Tyr Ser Pro Phe Lys
                245                 250                 255

gcc gtg tgg gac tgg ctc atc ctg ctg ctg gtc atc tac acg gct gtc   816
Ala Val Trp Asp Trp Leu Ile Leu Leu Leu Val Ile Tyr Thr Ala Val
            260                 265                 270

ttc acg ccc tac tca gcc gcc ttc ctg ctc agc gac cag gac gaa tca   864
Phe Thr Pro Tyr Ser Ala Ala Phe Leu Leu Ser Asp Gln Asp Glu Ser
        275                 280                 285

cgg cgt ggg gcc tgc agc tat acc tgc agt ccc ctc act gtg gtg gat   912
Arg Arg Gly Ala Cys Ser Tyr Thr Cys Ser Pro Leu Thr Val Val Asp
    290                 295                 300

ctc atc gtg gac atc atg ttc gtc gtg gac atc gtc atc aac ttc cgc   960
Leu Ile Val Asp Ile Met Phe Val Val Asp Ile Val Ile Asn Phe Arg
305                 310                 315                 320

acc acc tat gtc aac acc aat gat gag gtg gtc agc cac ccc cgc cgc   1008
Thr Thr Tyr Val Asn Thr Asn Asp Glu Val Val Ser His Pro Arg Arg
                325                 330                 335

atc gcc gtc cac tac ttc aag ggc tgg ttc ctc att gac atg gtg gcc   1056
Ile Ala Val His Tyr Phe Lys Gly Trp Phe Leu Ile Asp Met Val Ala
            340                 345                 350

gcc atc cct ttc gac ctc ctg atc ttc cgc act ggc tcc gat gag acc   1104
Ala Ile Pro Phe Asp Leu Leu Ile Phe Arg Thr Gly Ser Asp Glu Thr
        355                 360                 365

aca acc ctg att ggg cta ttg aag aca gcg cgg ctg ctg cgg ctg gtg   1152
Thr Thr Leu Ile Gly Leu Leu Lys Thr Ala Arg Leu Leu Arg Leu Val
    370                 375                 380

cgc gta gca cgg aag ctg gac cgc tac tct gag tat ggg gcg gct gtg   1200
Arg Val Ala Arg Lys Leu Asp Arg Tyr Ser Glu Tyr Gly Ala Ala Val
385                 390                 395                 400

ctc ttc ttg ctc atg tgc acc ttc gcg ctc ata gcg cac tgg ctg gcc   1248
Leu Phe Leu Leu Met Cys Thr Phe Ala Leu Ile Ala His Trp Leu Ala
                405                 410                 415

tgc atc tgg tac gcc atc ggc aat gtg gag cgg ccc tac cta gaa cac   1296
```

```
Cys Ile Trp Tyr Ala Ile Gly Asn Val Glu Arg Pro Tyr Leu Glu His
            420                 425                 430

aag atc ggc tgg ctg gac agc ctg ggt gtg cag ctt ggc aag cgc tac   1344
Lys Ile Gly Trp Leu Asp Ser Leu Gly Val Gln Leu Gly Lys Arg Tyr
        435                 440                 445

aac ggc agc gac cca gcc tcg ggc ccc tcg gtg cag gac aag tat gtc   1392
Asn Gly Ser Asp Pro Ala Ser Gly Pro Ser Val Gln Asp Lys Tyr Val
    450                 455                 460

aca gcc ctc tac ttc acc ttc agc agc ctc acc agc gtg ggc ttc ggc   1440
Thr Ala Leu Tyr Phe Thr Phe Ser Ser Leu Thr Ser Val Gly Phe Gly
465                 470                 475                 480

aat gtc tcg ccc aac acc aac tcc gag aag gtc ttc tcc atc tgc gtc   1488
Asn Val Ser Pro Asn Thr Asn Ser Glu Lys Val Phe Ser Ile Cys Val
                485                 490                 495

atg ctc atc ggc tcc ctg atg tac gcc agc atc ttc ggg aac gtg tcc   1536
Met Leu Ile Gly Ser Leu Met Tyr Ala Ser Ile Phe Gly Asn Val Ser
            500                 505                 510

gcg atc atc cag cgc ctg tac tcg ggc acc gcg cgc tac cac acg cag   1584
Ala Ile Ile Gln Arg Leu Tyr Ser Gly Thr Ala Arg Tyr His Thr Gln
        515                 520                 525

atg ctg cgt gtc aag gag ttc atc cgc ttc cac cag atc ccc aac cca   1632
Met Leu Arg Val Lys Glu Phe Ile Arg Phe His Gln Ile Pro Asn Pro
    530                 535                 540

ctg cgc cag cgc ctg gag gag tat ttc cag cac gcc tgg tcc tac acc   1680
Leu Arg Gln Arg Leu Glu Glu Tyr Phe Gln His Ala Trp Ser Tyr Thr
545                 550                 555                 560

aat ggc att gac atg aac gcg gtg ctg aag ggc ttc ccc gag tgc ctg   1728
Asn Gly Ile Asp Met Asn Ala Val Leu Lys Gly Phe Pro Glu Cys Leu
                565                 570                 575

cag gct gac atc tgc ctg cac ctg cac cgc gca ctg ctg cag cac tgc   1776
Gln Ala Asp Ile Cys Leu His Leu His Arg Ala Leu Leu Gln His Cys
            580                 585                 590

cca gct ttc agc ggc gcc ggc aag ggc tgc ctg cgc gcg cta gcc gtc   1824
Pro Ala Phe Ser Gly Ala Gly Lys Gly Cys Leu Arg Ala Leu Ala Val
        595                 600                 605

aag ttc aag acc acc cac gcg ccg cct ggg gac acg ctg gtg cac ctc   1872
Lys Phe Lys Thr Thr His Ala Pro Pro Gly Asp Thr Leu Val His Leu
    610                 615                 620

ggc gac gtg ctc tcc acc ctc tac ttc atc tcc cga ggc tcc atc gag   1920
Gly Asp Val Leu Ser Thr Leu Tyr Phe Ile Ser Arg Gly Ser Ile Glu
625                 630                 635                 640

atc ctg cgc gac gac gtg gtc gtg gcc atc cta gga aag aat gac atc   1968
Ile Leu Arg Asp Asp Val Val Val Ala Ile Leu Gly Lys Asn Asp Ile
                645                 650                 655

ttt ggg gaa ccc gtc agc ctc cat gcc cag cca ggc aag tcc agt gca   2016
Phe Gly Glu Pro Val Ser Leu His Ala Gln Pro Gly Lys Ser Ser Ala
            660                 665                 670
```

```
gac gtg cgg gct ctg acc tac tgc gac ctg cac aag atc cag cgg gca   2064
Asp Val Arg Ala Leu Thr Tyr Cys Asp Leu His Lys Ile Gln Arg Ala
        675                 680                 685

gat ctg ctg gag gtg ctg gac atg tac ccg gcc ttt gcg gag agc ttc   2112
Asp Leu Leu Glu Val Leu Asp Met Tyr Pro Ala Phe Ala Glu Ser Phe
    690                 695                 700

tgg agt aag ctg gag gtc acc ttc aac ctg cgg gac gca gcc ggg ggt   2160
Trp Ser Lys Leu Glu Val Thr Phe Asn Leu Arg Asp Ala Ala Gly Gly
705                 710                 715                 720

ctc cac tca tcc ccc cga cag gct cct ggc agc caa gac cac caa ggt   2208
Leu His Ser Ser Pro Arg Gln Ala Pro Gly Ser Gln Asp His Gln Gly
                725                 730                 735

ttc ttt ctc agt gac aac cag tca gat gca gcc cct ccc ctg agc atc   2256
Phe Phe Leu Ser Asp Asn Gln Ser Asp Ala Ala Pro Pro Leu Ser Ile
            740                 745                 750

tca gat gca tct ggc ctc tgg cct gag cta ctg cag gaa atg ccc cca   2304
Ser Asp Ala Ser Gly Leu Trp Pro Glu Leu Leu Gln Glu Met Pro Pro
        755                 760                 765

agg cac agc ccc caa agc cct cag gaa gac cca gat tgc tgg cct ctg   2352
Arg His Ser Pro Gln Ser Pro Gln Glu Asp Pro Asp Cys Trp Pro Leu
    770                 775                 780

aag ctg ggc tcc agg cta gag cag ctc cag gcc cag atg aac agg ctg   2400
Lys Leu Gly Ser Arg Leu Glu Gln Leu Gln Ala Gln Met Asn Arg Leu
785                 790                 795                 800

gag tcc cgc gtg tcc tca gac ctc agc cgc atc ttg cag ctc ctc cag   2448
Glu Ser Arg Val Ser Ser Asp Leu Ser Arg Ile Leu Gln Leu Leu Gln
                805                 810                 815

aag ccc atg ccc cag ggc cac gcc agc tac att ctg gaa gcc cct gcc   2496
Lys Pro Met Pro Gln Gly His Ala Ser Tyr Ile Leu Glu Ala Pro Ala
            820                 825                 830


tcc aat gac ctg gcc ttg gtt cct ata gcc tcg gag acg acg agt cca   2544
Ser Asn Asp Leu Ala Leu Val Pro Ile Ala Ser Glu Thr Thr Ser Pro
        835                 840                 845

ggg ccc agg ctg ccc cag ggc ttt ctg cct cct gca cag acc cca agc   2592
Gly Pro Arg Leu Pro Gln Gly Phe Leu Pro Pro Ala Gln Thr Pro Ser
    850                 855                 860

tat gga gac ttg gat gac tgt agt cca aag cac agg aac tcc tcc ccc   2640
Tyr Gly Asp Leu Asp Asp Cys Ser Pro Lys His Arg Asn Ser Ser Pro
865                 870                 875                 880

agg atg cct cac ctg gct gtg gca acg gac aaa act ctg gca cca tcc   2688
Arg Met Pro His Leu Ala Val Ala Thr Asp Lys Thr Leu Ala Pro Ser
                885                 890                 895

tca gaa cag gaa cag cct gag ggg ctc tgg cca ccc cta gcc tca cct   2736
Ser Glu Gln Glu Gln Pro Glu Gly Leu Trp Pro Pro Leu Ala Ser Pro
            900                 905                 910
```

```
cta cat ccc ctg gaa gta caa gga ctc atc tgt ggt ccc tgc ttc tcc   2784
Leu His Pro Leu Glu Val Gln Gly Leu Ile Cys Gly Pro Cys Phe Ser
        915                 920                 925

tcc ctc cct gaa cac ctt ggc tct gtt ccc aag cag ctg gac ttc cag   2832
Ser Leu Pro Glu His Leu Gly Ser Val Pro Lys Gln Leu Asp Phe Gln
    930                 935                 940

aga cat ggc tca gat cct gga ttt gca ggg agt tgg ggc cac tga       2877
Arg His Gly Ser Asp Pro Gly Phe Ala Gly Ser Trp Gly His  *
945                 950                 955
```

<210> 21
<211> 958
<212> PRT
<213> Homo Sapiens

<400> 21

```
Met Pro Val Arg Arg Gly His Val Ala Pro Gln Asn Thr Tyr Leu Asp
 1               5                  10                  15
Thr Ile Ile Arg Lys Phe Glu Gly Gln Ser Arg Lys Phe Leu Ile Ala
            20                  25                  30
Asn Ala Gln Met Glu Asn Cys Ala Ile Ile Tyr Cys Asn Asp Gly Phe
        35                  40                  45
Cys Glu Leu Phe Gly Tyr Ser Arg Val Glu Val Met Gln Gln Pro Cys
    50                  55                  60
Thr Cys Asp Phe Leu Thr Gly Pro Asn Thr Pro Ser Ser Ala Val Ser
65                  70                  75                  80
Arg Leu Ala Gln Ala Leu Leu Gly Ala Glu Glu Cys Lys Val Asp Ile
                85                  90                  95
Leu Tyr Tyr Arg Lys Asp Ala Ser Ser Phe Arg Cys Leu Val Asp Val
            100                 105                 110
Val Pro Val Lys Asn Glu Asp Gly Ala Val Ile Met Phe Ile Leu Asn
        115                 120                 125
Phe Glu Asp Leu Ala Gln Leu Leu Ala Lys Cys Ser Ser Arg Ser Leu
    130                 135                 140
Ser Gln Arg Leu Leu Ser Gln Ser Phe Leu Gly Ser Glu Gly Ser His
145                 150                 155                 160
Gly Arg Pro Gly Gly Pro Gly Pro Gly Thr Gly Arg Gly Lys Tyr Arg
                165                 170                 175
Thr Ile Ser Gln Ile Pro Gln Phe Thr Leu Asn Phe Val Glu Phe Asn
            180                 185                 190
Leu Glu Lys His Arg Ser Ser Ser Thr Thr Glu Ile Glu Ile Ile Ala
        195                 200                 205
Pro His Lys Val Val Glu Arg Thr Gln Asn Val Thr Glu Lys Val Thr
    210                 215                 220
Gln Val Leu Ser Leu Gly Ala Asp Val Leu Pro Glu Tyr Lys Leu Gln
225                 230                 235                 240
Ala Pro Arg Ile His Arg Trp Thr Ile Leu His Tyr Ser Pro Phe Lys
                245                 250                 255
Ala Val Trp Asp Trp Leu Ile Leu Leu Leu Val Ile Tyr Thr Ala Val
            260                 265                 270
Phe Thr Pro Tyr Ser Ala Ala Phe Leu Leu Ser Asp Gln Asp Glu Ser
        275                 280                 285
Arg Arg Gly Ala Cys Ser Tyr Thr Cys Ser Pro Leu Thr Val Val Asp
    290                 295                 300
Leu Ile Val Asp Ile Met Phe Val Val Asp Ile Val Ile Asn Phe Arg
305                 310                 315                 320
Thr Thr Tyr Val Asn Thr Asn Asp Glu Val Val Ser His Pro Arg Arg
                325                 330                 335
```

```
Ile Ala Val His Tyr Phe Lys Gly Trp Phe Leu Ile Asp Met Val Ala
            340                 345                 350
Ala Ile Pro Phe Asp Leu Leu Ile Phe Arg Thr Gly Ser Asp Glu Thr
        355                 360                 365
Thr Thr Leu Ile Gly Leu Leu Lys Thr Ala Arg Leu Leu Arg Leu Val
    370                 375                 380
Arg Val Ala Arg Lys Leu Asp Arg Tyr Ser Glu Tyr Gly Ala Ala Val
385                 390                 395                 400
Leu Phe Leu Leu Met Cys Thr Phe Ala Leu Ile Ala His Trp Leu Ala
                405                 410                 415
Cys Ile Trp Tyr Ala Ile Gly Asn Val Glu Arg Pro Tyr Leu Glu His
            420                 425                 430
Lys Ile Gly Trp Leu Asp Ser Leu Gly Val Gln Leu Gly Lys Arg Tyr
        435                 440                 445
Asn Gly Ser Asp Pro Ala Ser Gly Pro Ser Val Gln Asp Lys Tyr Val
    450                 455                 460
Thr Ala Leu Tyr Phe Thr Phe Ser Ser Leu Thr Ser Val Gly Phe Gly
465                 470                 475                 480
Asn Val Ser Pro Asn Thr Asn Ser Glu Lys Val Phe Ser Ile Cys Val
                485                 490                 495
Met Leu Ile Gly Ser Leu Met Tyr Ala Ser Ile Phe Gly Asn Val Ser
            500                 505                 510
Ala Ile Ile Gln Arg Leu Tyr Ser Gly Thr Ala Arg Tyr His Thr Gln
        515                 520                 525
Met Leu Arg Val Lys Glu Phe Ile Arg Phe His Gln Ile Pro Asn Pro
    530                 535                 540
Leu Arg Gln Arg Leu Glu Glu Tyr Phe Gln His Ala Trp Ser Tyr Thr
545                 550                 555                 560
Asn Gly Ile Asp Met Asn Ala Val Leu Lys Gly Phe Pro Glu Cys Leu
                565                 570                 575
Gln Ala Asp Ile Cys Leu His Leu His Arg Ala Leu Leu Gln His Cys
            580                 585                 590
Pro Ala Phe Ser Gly Ala Gly Lys Gly Cys Leu Arg Ala Leu Ala Val
        595                 600                 605
Lys Phe Lys Thr Thr His Ala Pro Pro Gly Asp Thr Leu Val His Leu
    610                 615                 620
Gly Asp Val Leu Ser Thr Leu Tyr Phe Ile Ser Arg Gly Ser Ile Glu
625                 630                 635                 640
Ile Leu Arg Asp Asp Val Val Val Ala Ile Leu Gly Lys Asn Asp Ile
                645                 650                 655
Phe Gly Glu Pro Val Ser Leu His Ala Gln Pro Gly Lys Ser Ser Ala
            660                 665                 670
Asp Val Arg Ala Leu Thr Tyr Cys Asp Leu His Lys Ile Gln Arg Ala
        675                 680                 685
Asp Leu Leu Glu Val Leu Asp Met Tyr Pro Ala Phe Ala Glu Ser Phe
    690                 695                 700
Trp Ser Lys Leu Glu Val Thr Phe Asn Leu Arg Asp Ala Ala Gly Gly
705                 710                 715                 720
Leu His Ser Ser Pro Arg Gln Ala Pro Gly Ser Gln Asp His Gln Gly
                725                 730                 735
Phe Phe Leu Ser Asp Asn Gln Ser Asp Ala Ala Pro Pro Leu Ser Ile
            740                 745                 750
Ser Asp Ala Ser Gly Leu Trp Pro Glu Leu Leu Gln Glu Met Pro Pro
        755                 760                 765
Arg His Ser Pro Gln Ser Pro Gln Glu Asp Pro Asp Cys Trp Pro Leu
    770                 775                 780
Lys Leu Gly Ser Arg Leu Glu Gln Leu Gln Ala Gln Met Asn Arg Leu
785                 790                 795                 800
Glu Ser Arg Val Ser Ser Asp Leu Ser Arg Ile Leu Gln Leu Leu Gln
                805                 810                 815
Lys Pro Met Pro Gln Gly His Ala Ser Tyr Ile Leu Glu Ala Pro Ala
            820                 825                 830
```

```
Ser Asn Asp Leu Ala Leu Val Pro Ile Ala Ser Glu Thr Thr Ser Pro
        835                 840                 845
Gly Pro Arg Leu Pro Gln Gly Phe Leu Pro Pro Ala Gln Thr Pro Ser
    850                 855                 860
Tyr Gly Asp Leu Asp Asp Cys Ser Pro Lys His Arg Asn Ser Ser Pro
865                 870                 875                 880
Arg Met Pro His Leu Ala Val Ala Thr Asp Lys Thr Leu Ala Pro Ser
                885                 890                 895
Ser Glu Gln Glu Gln Pro Glu Gly Leu Trp Pro Pro Leu Ala Ser Pro
            900                 905                 910
Leu His Pro Leu Glu Val Gln Gly Leu Ile Cys Gly Pro Cys Phe Ser
        915                 920                 925

Ser Leu Pro Glu His Leu Gly Ser Val Pro Lys Gln Leu Asp Phe Gln
    930                 935                 940
Arg His Gly Ser Asp Pro Gly Phe Ala Gly Ser Trp Gly His
945                 950                 955
```