US 20090226449A1

(19) United States

(12) Patent Application Publication

Glucksmann et al.

(10) Pub. No.: US 2009/0226449 A1

(43) Pub. Date: Sep. 10, 2009

(54) **NOVEL 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921, AND 33751 MOLECULES AND USES THEREFOR**

(75) Inventors: **Maria A. Glucksmann**, Lexington, MA (US); **Rory A.J. Curtis**, Ashland, MA (US); **Jose M. Lora**, Mountain View, CA (US); **Katherine M. Galvin**, Jamaica Plain, MA (US); **Inmaculada Silos-Santiago**, Del Mar, CA (US)

Correspondence Address:
MILLENNIUM PHARMACEUTICALS, INC.
40 Landsdowne Street
CAMBRIDGE, MA 02139 (US)

(73) Assignee: **Millennium Pharmaceuticals, Inc.**, Cambridge, MA (US)

(21) Appl. No.: **12/381,748**

(22) Filed: **Mar. 16, 2009**

Related U.S. Application Data

(63) Continuation of application No. 10/391,399, filed on Mar. 18, 2003, which is a continuation-in-part of application No. 09/789,481, filed on Feb. 20, 2001, now abandoned, which is a continuation-in-part of application No. 09/634,669, filed on Aug. 8, 2000, now abandoned, which is a continuation-in-part of application No. 09/583,373, filed on May 31, 2000, now abandoned, which is a continuation-in-part of application No. 09/510,706, filed on Feb. 22, 2000, now abandoned, said application No. 10/391,399 is a continuation-in-part of application No. 10/309,804, filed on Dec. 4, 2002, now abandoned, said application No. 10/391,399 is a continuation-in-part of application No. 10/094,214, filed on Mar. 8, 2002, now abandoned, said application No. 10/391,399 is a continuation-in-part of application No. 09/828,035, filed on Apr. 6, 2001, now abandoned, said application No. 10/391,399 is a continuation-in-part of application No. 09/891,762, filed on Jun. 26, 2001, now abandoned, said application No. 10/391,399 is a continuation-in-part of application No. 10/245,121, filed on Sep. 17, 2002, now abandoned, said application No. 10/391,399 is a continuation-in-part of application No. 10/095,139, filed on Mar. 11, 2002, now abandoned, said application No. 10/391,399 is a continuation-in-part of application No. 09/957,683, filed on Sep. 19, 2001, now abandoned, said application No. 10/391,399 is a continuation-in-part of application No. 09/942,447, filed on Aug. 29, 2001, now abandoned, said application No. 10/391,399 is a continuation-in-part of application No. 10/062,937, filed on Jan. 31, 2002, now abandoned, said application No. 10/391,399 is a continuation-in-part of application No. 10/255,532, filed on Sep. 26, 2002, now abandoned.

(60) Provisional application No. 60/336,936, filed on Dec. 4, 2001, provisional application No. 60/275,078, filed on Mar. 12, 2001, provisional application No. 60/195,734, filed on Apr. 7, 2000, provisional application No. 60/214,176, filed on Jun. 26, 2000, provisional application No. 60/322,983, filed on Sep. 17, 2001, provisional application No. 60/275,172, filed on Mar. 12, 2001, provisional application No. 60/233,537, filed on Sep. 19, 2000, provisional application No. 60/229,036, filed on Aug. 31, 2000, provisional application No. 60/267,076, filed on Feb. 1, 2001, provisional application No. 60/325,854, filed on Sep. 27, 2001.

Publication Classification

(51) **Int. Cl.**

A61K 39/395	(2006.01)
C12N 15/12	(2006.01)
C12N 5/10	(2006.01)
C07K 14/47	(2006.01)
C07K 16/18	(2006.01)
C12P 21/02	(2006.01)
C12Q 1/68	(2006.01)
G01N 33/566	(2006.01)
A61K 38/17	(2006.01)
A61K 31/7105	(2006.01)
A61K 38/48	(2006.01)
A61P 35/00	(2006.01)
A61P 43/00	(2006.01)

(52) **U.S. Cl.** ..... **424/139.1**; 536/23.5; 435/325; 530/350; 530/387.9; 435/69.1; 435/6; 436/501; 435/375; 514/12; 514/44 A; 514/44 R; 424/94.63

(57) **ABSTRACT**

The invention provides isolated nucleic acids molecules, designated 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 and 33751 nucleic acid molecules. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 and 33751 nucleic acid molecules, host cells into which the expression vectors have been introduced, and non-human transgenic animals in which a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene has been introduced or disrupted. The invention still further provides isolated 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins, fusion proteins, antigenic peptides and anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibodies. Diagnostic and therapeutic methods utilizing compositions of the invention are also provided.

NOVEL 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921, AND 33751 MOLECULES AND USES THEREFOR

RELATED APPLICATIONS

[0001] The present application is a continuation of U.S. patent application Ser. No. 10/391,399, filed Mar. 18, 2003 (pending), which is a continuation-in-part of U.S. patent application Ser. No. 09/789,481, filed Feb. 20, 2001 (abandoned), which is a continuation-in-part of U.S. patent application Ser. No. 09/634,669, filed Aug. 8, 2000 (abandoned), which is a continuation-in-part of U.S. patent application Ser. No. 09/583,373, filed May 31, 2000 (abandoned), which is a continuation-in-part of U.S. patent application Ser. No. 09/510,706, filed Feb. 22, 2000 (abandoned). U.S. patent application Ser. No. 10/391,399 is also a continuation-in-part of U.S. patent application Ser. No. 10/309,804, filed Dec. 4, 2002 (abandoned), which claims the benefit of U.S. Provisional Application Ser. No. 60/336,936, filed Dec. 4, 2001 (abandoned). U.S. patent application Ser. No. 10/391,399 is also a continuation-in-part of U.S. patent application Ser. No. 10/094,214, filed Mar. 8, 2002 (abandoned), which claims the benefit of U.S. Provisional Application Ser. No. 60/275,078, filed Mar. 12, 2001 (abandoned). U.S. patent application Ser. No. 10/391,399 is also a continuation-in-part of U.S. patent application Ser. No. 09/828,035, filed Apr. 6, 2001 (abandoned), which claims the benefit of U.S. Provisional Application Ser. No. 60/195,734, filed Apr. 7, 2000 (abandoned). U.S. patent application Ser. No. 10/391,399 is also a continuation-in-part of U.S. patent application Ser. No. 09/891,762, filed Jun. 26, 2001 (abandoned), which claims the benefit of U.S. Provisional Application Ser. No. 60/214,176, filed Jun. 26, 2000 (abandoned). U.S. patent application Ser. No. 10/391,399 is also a continuation-in-part of U.S. patent application Ser. No. 10/245,121, filed Sep. 17, 2002 (abandoned), which claims the benefit of U.S. Provisional Application Ser. No. 60/322,983, filed Sep. 17, 2001 (abandoned). U.S. patent application Ser. No. 10/391,399 is also a continuation-in-part of U.S. patent application Ser. No. 10/095,139, filed Mar. 11, 2002 (abandoned), which claims the benefit of U.S. Provisional Application Ser. No. 60/275,172, filed Mar. 12, 2001 (abandoned). U.S. patent application Ser. No. 10/391,399 is also a continuation-in-part of U.S. patent application Ser. No. 09/957,683, filed Sep. 19, 2001 (abandoned), which claims the benefit of U.S. Provisional Application Ser. No. 60/233,537, filed on Sep. 19, 2000 (abandoned). U.S. patent application Ser. No. 10/391,399 is also a continuation-in-part of U.S. patent application Ser. No. 09/942,447, filed Aug. 29, 2001 (abandoned), which claims the benefit of U.S. Provisional Application Ser. No. 60/229,036, filed Aug. 31, 2000 (abandoned). U.S. patent application Ser. No. 10/391,399 is also a continuation-in-part of U.S. patent application Ser. No. 10/062,937, filed Jan. 31, 2002 (abandoned), which claims the benefit of U.S. Provisional Application Ser. No. 60/267,076, filed Feb. 1, 2001 (abandoned). U.S. patent application Ser. No. 10/391,399 is also a continuation-in-part of U.S. patent application Ser. No. 10/255,532, filed Sep. 26, 2002 (abandoned), which claims the benefit of U.S. Provisional Application Ser. No. 60/325,854, filed Sep. 27, 2001 (abandoned). The entire contents of each of the above-referenced patent applications are incorporated herein by this reference.

BACKGROUND OF THE INVENTION

[0002] Many intercellular and intracellular events depend on the regulation of the concentration of certain ions such as calcium, sodium, potassium, and chloride within the cell. Ion channels have been found which regulate the flow of anions or cations through a membrane based on voltage (voltage-gated channels), pH (mechanically gated channels), phosphorylation, and ligand binding (ligand-gated channels). Typically, an ion channel consists of multiple transmembrane domains which form a channel through which the ions pass from one side of the membrane to the other. These ion channels may play important roles in how a cell responds to hormones or neurotransmitters with increased activity of enzymes such as phospholipase C and a subsequent rise in the concentration of intracellular free calcium (Ca^{+2}). The increase in intracellular calcium concentration may occur as a result of the release of calcium from intracellular stores as well as an influx of calcium through the plasma membrane.

[0003] Examples of ion channels include, for example, calcium channel, calcium/sodium antiporters, potassium channel, organic ion transporter, and choline transporters. Such ion channels have the ability, for example: 1) to modulate membrane excitability; 2) to influence the resting potential of membranes; 3) to modulate wave forms and frequencies of action potentials; 4) to modulate thresholds of excitation; 5) to modulate neurite outgrowth and synaptogenesis; 6) to modulate signal transduction, 7) to bind a second messenger; 8) to bind diacylglycerol; 9) to regulate the flow of cations through a membrane; 10) to transport a substrate or target molecule, e.g., an ion (e.g., a calcium ion) across a membrane; 11) to transport a second substrate or target molecule, e.g., another ion (e.g., a sodium ion) across a membrane; 12) to transport a third substrate or target molecule, e.g., another ion (e.g., a potassium ion) across a membrane; 13) to interact with and/or modulate the activity of a second non-transporter protein; 14) to modulate cellular signaling and/or gene transcription (e.g., either directly or indirectly; 15) to interact with a non-TWIK protein molecule; 16) to activate a TWIK-dependent signal transduction pathway; 17) to modulate the release of neurotransmitters; 18) to protect cells and/or tissues from organic ions; 19) to modulate intracellular Ca^{2+} concentration; 20) to bind a ligand, e.g., L-glutamate, and/or glycine; 21) to modulate (e.g., promote, catalyze, regulate, initiate, facilitate or inhibit) the manufacture of choline metabolites and/or compounds of which choline is a component or precursor, e.g., phospholipids (e.g., phosphatidylcholine (lecithin), sphingomyelin, sphingophosphorylcholine, and platelet activating factor), acetylcholine, very low density lipoproteins (VLDLs), and betaine, e.g., by transporting choline into or out of cells; 22) to modulate (e.g., promote, catalyze, regulate, initiate, facilitate or inhibit) transport of choline, its metabolites, and/or compounds of which choline is a component or precursor across membranes (e.g., plasma membranes), e.g., from an extracellular medium into a cell, or vice versa; 23) to modulate (e.g., promote, catalyze, regulate, initiate, facilitate or inhibit) transport of choline, its metabolites, and/or compounds of which choline is a component or precursor across barriers between tissues (e.g., the blood-brain barrier); as well as many others. Accordingly, there exists a need to identify additional ion channels, for example, for use as disease markers and as targets for identifying various therapeutic modulators.

SUMMARY OF THE INVENTION

[0004] The present invention is based, at least in part, on the discovery of novel nucleic acid molecules and proteins encoded by such nucleic acid molecules, referred to herein as

"18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751". The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid and protein molecules of the present invention are useful as modulating agents in regulating a variety of cellular processes, e.g., including but not limited to, modulating cellular response to hormones or neurotransmitters. In particular, these nucleic acid molecules will be advantageous in the regulation of any cellular function, uncontrolled proliferation and differentiation, such as in cases of pain. Accordingly, in one aspect, this invention provides isolated nucleic acid molecules encoding 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-encoding nucleic acids.

[0005] The nucleotide sequence of the cDNA encoding 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751, and the amino acid sequence of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptides are depicted in Table 1.

TABLE 1

Sequences of the invention

Gene Name	cDNA	Protein	Coding Region
18607	SEQ ID NO: 1	SEQ ID NO: 2	SEQ ID NO: 3
15603	SEQ ID NO: 5	SEQ ID NO: 6	SEQ ID NO: 7
69318	SEQ ID NO: 13	SEQ ID NO: 14	SEQ ID NO: 15
12303	SEQ ID NO: 18	SEQ ID NO: 19	SEQ ID NO: 20
48000	SEQ ID NO: 21	SEQ ID NO: 22	SEQ ID NO: 23
52920	SEQ ID NO: 24	SEQ ID NO: 25	SEQ ID NO: 26
5433	SEQ ID NO: 27	SEQ ID NO: 28	SEQ ID NO: 29
38554	SEQ ID NO: 32 and 53	SEQ ID NO: 33 and 54	SEQ ID NO: 34 and 55
57301	SEQ ID NO: 35	SEQ ID NO: 36	SEQ ID NO: 37
58324	SEQ ID NO: 38	SEQ ID NO: 39	SEQ ID NO: 40
55063	SEQ ID NO: 56	SEQ ID NO: 57	SEQ ID NO: 58
52991	SEQ ID NO: 60	SEQ ID NO: 61	SEQ ID NO: 62
59914	SEQ ID NO: 69	SEQ ID NO: 70	SEQ ID NO: 71
59921	SEQ ID NO: 72	SEQ ID NO: 73	SEQ ID NO: 74
33751	SEQ ID NO: 81	SEQ ID NO: 82	SEQ ID NO: 83

[0006] Accordingly, in one aspect, the invention features a nucleic acid molecule which encodes a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or polypeptide, e.g., a biologically active portion of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. In a preferred embodiment, the isolated nucleic acid molecule encodes a polypeptide having the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82. In other embodiments, the invention provides isolated 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid molecules having the nucleotide sequence shown in SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83. In still other embodiments, the invention provides nucleic acid molecules that are substantially identical (e.g., naturally occurring allelic variants) to the nucleotide sequence shown in SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83. In other embodiments, the invention provides a nucleic acid molecule which hybridizes under a stringent hybridization condition as described herein to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, wherein the nucleic acid encodes a full length 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or an active fragment thereof.

[0007] In a related aspect, the invention further provides nucleic acid constructs which include a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid molecule described herein. In certain embodiments, the nucleic acid molecules of the invention are operatively linked to native or heterologous regulatory sequences. Also included are vectors and host cells containing the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid molecules of the invention e.g., vectors and host cells suitable for producing polypeptides.

[0008] In another related aspect, the invention provides nucleic acid fragments suitable as primers or hybridization probes for the detection of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-encoding nucleic acids.

[0009] In still another related aspect, isolated nucleic acid molecules that are antisense to a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 encoding nucleic acid molecule are provided.

[0010] In another aspect, the invention features 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptides, and biologically active or antigenic fragments thereof that are useful, e.g., as reagents or targets in assays applicable to treatment and diagnosis of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disorders. In another embodiment, the invention provides 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptides having a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity.

[0011] In other embodiments, the invention provides 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptides, e.g., a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide having the amino acid sequence shown in SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82; an amino acid sequence that is substantially identical to the amino acid sequence shown in SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82; or an amino acid sequence encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under a stringent hybridization condition as described herein to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, wherein the nucleic acid encodes a full length 18607,

15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or an active fragment thereof.

[0012] In a related aspect, the invention further provides nucleic acid constructs which include a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid molecule described herein.

[0013] In a related aspect, the invention provides 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptides or fragments operatively linked to non-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptides to form fusion proteins.

[0014] In another aspect, the invention features antibodies and antigen-binding fragments thereof, that react with, or more preferably specifically or selectively bind 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptides.

[0015] In another aspect, the invention provides methods of screening for compounds that modulate the expression or activity of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptides or nucleic acids.

[0016] In still another aspect, the invention provides a process for modulating 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide or nucleic acid expression or activity, e.g., using the compounds identified in the screens described herein. In certain embodiments, the methods involve treatment of conditions related to aberrant activity or expression of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptides or nucleic acids, such as conditions or disorders involving aberrant or deficient 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression. Examples of such disorders include, but are not limited to cellular proliferative and/or differentiative disorders, brain disorders, blood vessel disorders, platelet disorders, breast disorders, colon disorders, kidney disorders, lung disorders, ovarian disorders, prostate disorders, pancreatic disorders, skeletal muscle disorders, testicular disorders, eye disorders, hormonal disorders, disorders associated with bone metabolism, immune e.g., inflammatory, disorders, cardiovascular disorders, endothelial cell disorders, liver disorders, viral diseases, pain, metabolic disorders, neurological disorders, neurodegenerative disorders or angiogenic disorders.

[0017] The invention also provides assays for determining the activity of or the presence or absence of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptides or nucleic acid molecules in a biological sample, including for disease diagnosis.

[0018] In a further aspect, the invention provides assays for determining the presence or absence of a genetic alteration in a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide or nucleic acid molecule, including for disease diagnosis.

[0019] In another aspect, the invention features a two dimensional array having a plurality of addresses, each address of the plurality being positionally distinguishable from each other address of the plurality, and each address of the plurality having a unique capture probe, e.g., a nucleic acid or peptide sequence. At least one address of the plurality has a capture probe that recognizes a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecule. In one embodiment, the capture probe is a nucleic acid, e.g., a probe complementary to a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid sequence. In another embodiment, the capture probe is a polypeptide, e.g., an antibody specific for 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptides. Also featured is a method of analyzing a sample by contacting the sample to the aforementioned array and detecting binding of the sample to the array.

[0020] Other features and advantages of the invention will be apparent from the following detailed description, and from the claims.

DETAILED DESCRIPTION OF THE INVENTION

Human 18607 (TLCC)

[0021] The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as "TRP-like calcium channel", "TLCC" or "18607" nucleic acid and protein molecules, which are novel members of the calcium channel family. These novel molecules are capable of, for example, modulating a calcium channel mediated activity in a cell, e.g., a neuronal, muscle (e.g., cardiac muscle), or liver cell. The present invention is further based, at least in part, on the discovery that TLCC genes are up-regulated in stellate cells (the main effectors of liver fibrosis) as compared to their expression in normal hepatic cells, and, thus, may be associated with a hepatic disorder. Accordingly, the present invention further provides methods and compositions for the diagnosis and treatment of a hepatic disorder, including but not limited to, liver fibrosis, hepatitis, liver tumors, cirrhosis of the liver, hemochromatosis, liver parasite induced disorders, alpha-1 antitrypsin deficiency, and autoimmune hepatitis.

[0022] The human TLCC or 18607 sequence (SEQ ID NO:1), which is approximately 3900 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 3387 nucleotides (nucleotides 138-3524 of SEQ ID NO:1; nucleotides 1-3387 of SEQ ID NO:3), not including the terminal codon. The coding sequence encodes a 1129 amino acid protein (SEQ ID NO:2).

[0023] A BLASTN 2.0 search against the NRN database, using a score of 100 and a word length of 12 (Altschul et al. (1990) *J. Mol. Biol.* 215:403) of the nucleotide sequence of human TLCC revealed that human TLCC is 97% identical to human STS WI-30695, sequence tagged site (Accession Number G22461) over nucleotides 3874-3605. This search further revealed that human TLCC is homologous to human chromosome 11p15.5 PAC clone pDJ915f1 containing KvLQT1 gene, complete sequence (Accession Number AC003693).

[0024] A BLASTN 2.0 search against the dbEST database, using a score of 100 and a word length of 12 (Altschul et al. (1990) *J. Mol. Biol.* 215:403) of the nucleotide sequence of human TLCC revealed that human TLCC is 98% identical to nf99c01.s1 NCI_CGAP_Co3 *Homo sapiens* cDNA clone IMAGE:928032 (Accession Number AA551759) over nucleotides 3865-3369 of SEQ ID NO:1. This search further

revealed that human TLCC is 100% identical to tg78b06.x1Soares_NhHMPu_S1 *Homo sapiens* cDNA clone IMAGE:2114867 (Accession Number AI417040) over nucleotides 3866-3391 of SEQ ID NO:1. This search further revealed that human TLCC is 97% identical to nq58f08.s1 NCI_CGAP_Co9 *Homo sapiens* cDNA clone IMAGE: 1148103 (Accession Number AA633315) over nucleotides 3868-3428 of SEQ ID NO:1. This search further revealed that human TLCC is 98% identical to qp09f02.x1 NCI_CGAP_Kid5 *Homo sapiens* cDNA clone IMAGE:1917531 3', mRNA sequence (Accession Number AI344661) over nucleotides 3866-3437 of SEQ ID NO:1. This search further revealed that human TLCC is 97% identical to ah33h08.s1 Soares testis NHT *Homo sapiens* cDNA clone 1276383 3' (Accession Number AA694490) over nucleotides 3863-3418 of SEQ ID NO:1.

[0025] A BLASTN 2.0 search against the PATENT_2/gsnuc database, using a score of 100 and a wordlength of 12, of the nucleotide sequence of human TLCC revealed that human TLCC is 98% identical to human PS112 consensus DNA fragment from gene specific clones (Accession Number V26656) over nucleotides 1509-3900 of SEQ ID NO:1. This search further revealed that human TLCC is 99% identical to full-length cDNA sequence of prostate tumor clone J1-17 (Accession Number V61200) over nucleotides 2360-3881 of SEQ ID NO:1. This search further revealed that human TLCC is 99% identical to prostate tumour specific gene clone J1-17 (Accession Number V58585) over nucleotides 2360-3881 of SEQ ID NO:1. This search further revealed that human TLCC is 99% identical to human PS112 5'-EST DNA fragment (Accession Number V26657) over nucleotides 2614-3900 of SEQ ID NO:1. This search further revealed that human TLCC is 94% identical to 3' cDNA sequence of prostate tumor clone J1-17 (Accession Number V61142) over nucleotides 3204-3755 of SEQ ID NO:1. This search further found that human TLCC is 94% identical to 3' fragment of prostate tumour specific gene J1-17 (Accession Number V58485) over nucleotides 3204-3755 of SEQ ID NO:1. A CLUSTAL W (1.74) alignment of the human TLCC nucleotide sequence with the top hit in this search confirms the similarity of the sequences.

[0026] A BLASTN 2.0 search against the PATENT_2/Patent DbPreviewNuc database, using a score of 100 and a wordlength of 12, of the nucleotide sequence of human TLCC revealed that human TLCC is 99% identical to human nucleic acid (Accession Number AC31503 (WO99/46374)) over nucleotides 2339-3886 of SEQ ID NO:1, and 56% identical over nucleotides 3778-3895 of SEQ ID NO:1. This search further revealed that human TLCC is 99% identical to human nucleic acid (Accession Number AC31066 (WO99/46374)) over nucleotides 2621-3170 of SEQ ID NO:1. This search further revealed that human TLCC is 62% identical to 36 secreted proteins (Accession Number AC28066 (WO99/35158)) over nucleotides 2261-3173 of SEQ ID NO:1. This search further revealed that human TLCC is 64% identical to 36 secreted proteins (Accession Number AC28051 (WO99/35158)) over nucleotides 2421-3173 of SEQ ID NO:1.

[0027] A BLASTX 2.0 search against the NRP/protot database, using a wordlength of 3, a score of 100, and a BLOSUM62 matrix, of the translated nucleotide sequence of human TLCC revealed that human TLCC is 35% identical to the amino acid sequence of *C. elegans* hypothetical protein CET01H8.1, CEC05C12.3, CEF54D1.5 similar to trp and trp-like proteins [*Homo sapiens*] (Accession Number AB001535) over translated nucleic acid residues 147 to 2018 of SEQ ID NO:1, and 41% identical over translated nucleic acid residues 2205-3470 of SEQ ID NO:1. This search further found that human TLCC is 32% identical to the amino acid sequence of Accession Number Z83117, similarity with *Drosophila* transient-reporter-potential protein (Swiss Prot accession number P19334); cDNA EST EMBL: D27562 comes from this gene, cDNA EST yk219f12.5 comes from this gene [*Caenorhabditis elegans*] over translated nucleic acid residues 84-1418 of SEQ ID NO:1, 27% identical over translated nucleic acid residues 2190-3368 of SEQ ID NO:1, 30% identical over translated nucleic acid residues 1470-2063 of SEQ ID NO:1, 28% identical over translated nucleic acid residues 3076-3213 of SEQ ID NO:1, 46% identical over translated nucleic acid residues 1613-1651 of SEQ ID NO:1, and 32% identical over translated nucleic acid residues 3705-3839 of SEQ ID NO:1. This search further found that human TLCC is 33% identical to *Homo sapiens* melastatin I (Accession Number AF071787) over translated nucleic acid residues 2205-3401 of SEQ ID NO:1, 33% identical over translated nucleic acid residues 150-1142 of SEQ ID NO:1, 27% identical over translated nucleic acid residues 1548 to 2405 of SEQ ID NO:1, 48% identical over translated nucleic acid residues 1155-1298, 34% identical over translated nucleic acid residues 3801-3896 of SEQ ID NO:1, 30% identical over translated nucleic acid residues 1261-1380 of SEQ ID NO:1, and 36% identical over translated nucleic acid residues 2451-2516 of SEQ ID NO:1. This search further found that human TLCC is 31% identical to cDNA EST yk308e9.3 comes from this gene; cDNA EST yk308e9.5 comes from this gene; cDNA EST yk318f4.3 comes from this gene; cDNA EST yk318f4.5 comes from this gene; cDNA EST yk398a12.3 comes from this gene, cDNA EST yk398a12.5 comes from this gene (Accession Number Z68333) over translated nucleic acid residues 147-1328 of SEQ ID NO:1, is 23% identical over translated nucleic acid residues 2190-3422 of SEQ ID NO:1, is 31% identical over translated nucleic acid residues 1554-2099 of SEQ ID NO:1, is 34% identical over translated nucleic acid residues 1355-1468 of SEQ ID NO:1, and is 32% identical over translated nucleic acid residues 3225-3338 of SEQ ID NO:1. This search further found that human TLCC is 29% identical to similarity to Worm protein C05C12.3; cDNA EST yk224b10.3 comes from this gene; cDNA EST yk224b10.5 comes from this gene; cDNA EST yk301f12.3 comes from this gene; cDNA EST yk301f12.5 comes from this gene; cDNA EST yk405b7.3 comes from this gene over translated nucleic acid residues 147-2069 of SEQ ID NO:1, is 26% identical over translated nucleic acid residues 2193-2978 of SEQ ID NO:1, and is 34% identical over translated nucleic acid residues 2895-3257 of SEQ ID NO:1. This search further found that human TLCC is 34% identical to *Mus musculus* melastatin (Accession Number AF047714) over translated nucleic acid residues 150-1142 of SEQ ID NO:1, is 48% identical over translated nucleic acid residues 1155-1298 of SEQ ID NO:1, and is 36% identical over translated nucleic acid residues 2427-2516 of SEQ ID NO:1. A CLUSTAL W (1.74) alignment of the translated human TLCC sequence with the top three hits in this search confirms the similarity of the sequences.

[0028] A BLASTX 2.0 search against the PATENT_2/gsprot database, using a score of 100, a wordlength of 3 and a BLOSUM62 matrix, of the translated nucleotide sequence of human TLCC revealed that human TLCC is 95% identical to human PS112 protein sequence from gene-specific clones (Accession Number W54425) over translated nucleic acid

residues 1509-3524 of SEQ ID NO:1. This search further revealed that human TLCC is 100% identical to amino acid encoded by prostate tumour clone J1-17 (Accession Number W71868) over translated nucleic acid residues 2580-3524 of SEQ ID NO:1. This search further revealed that human TLCC is 100% identical to prostate tumour specific gene clone J1-17 protein (Accession Number W69384) over translated nucleic acid residues 2580-3524 of SEQ ID NO:1. This search further revealed that human TLCC is 34% identical to prostate-tumour derived antigen #4 (Accession Number Y00931) over translated nucleic acid residues 147-1310 of SEQ ID NO:1, 37% identical over translated nucleic acid residues 2457-3401 of SEQ ID NO:1, 36% identical over translated nucleic acid residues 1554-2018 of SEQ ID NO:1, 46% identical over translated nucleic acid residues 2196-2390 of SEQ ID NO:1, and 38% identical over translated nucleic acid residues 2931-2993 of SEQ ID NO:1. A ClustalW (1.74) alignment of the translated cDNA sequence of human TLCC with the top four hits of this search confirms the similarity of the sequences.

[0029] A search was performed against the Memsat database and correlated with an analysis of the hydrophilicity and surface probability of human TLCC, resulting in the identification of six transmembrane domains in the amino acid sequence of human TLCC (SEQ ID NO:2) at about residues 599-619, residues 690-712, residues 784-803, residues 811-831, residues 845-862, and residues 933-957.

[0030] A search was also performed against the Prosite database, and resulted in the identification of an N-glycosylation site at residues 143-146, at residues 205-208, and at residues 907-910 of SEQ ID NO:2.

[0031] A search was also performed against the ProDom database resulting in the identification of a "transmembrane calcium channel" domain in human TLCC (SEQ ID NO:2) at about residues 783-845. This search further identified significant sequence similarity between the amino acid sequence of human TLCC and human melastatin (Accession Number AAC80000). An alignment (using the GAP program in the GCG software package (Blosum 62 matrix), a gap weight of 12, and a length weight of 4) of the amino acid sequence of human TLCC with human melastatin (Accession Number AAC80000), revealed that human TLCC is 31.739% identical to human melastatin.

[0032] As used herein, a "calcium channel" includes a protein or polypeptide which is involved in receiving, conducting, and transmitting signals in an electrically excitable cell, e.g., a neuronal or muscle cell. Calcium channels are calcium ion selective, and can determine membrane excitability (the ability of, for example, a neuronal cell to respond to a stimulus and to convert it into a sensory impulse). Calcium channels can also influence the resting potential of membranes, waveforms and frequencies of action potentials, and thresholds of excitation. Calcium channels are typically expressed in electrically excitable cells, e.g., neuronal cells, and may form heteromultimeric structures (e.g., composed of more than one type of subunit). Calcium channels may also be found in nonexcitable cells (e.g., adipose cells or liver cells), where they may play a role in, e.g., signal transduction. Examples of calcium channels include the low-voltage-gated channels and the high-voltage-gated channels. Calcium channels are described in, for example, Davila et al. (1999) *Annals New York Academy of Sciences* 868:102-17 and McEnery, M. W. et al. (1998) *J. Bioenergetics and Biomembranes* 30(4): 409-418, the contents of which are incorporated herein by reference.

[0033] Calcium signaling has been implicated in the regulation of a variety of cellular responses, such as growth and differentiation. There are two general methods by which intracellular concentrations of calcium ions may be increased: calcium ions may be brought into the cell from the extracellular milieu through the use of specific channels in the cellular membrane, or calcium ions may be freed from intracellular stores, again being transported by specific membrane channels in the storage organelle. In the situation in which the intracellular stores of calcium have been depleted, a specific type of calcium channel, termed a 'capacitative calcium channel' or a 'store-operated calcium channel' (SOC), is activated in the plasma membrane to import calcium ions from the extracellular environment to the cytosol (for review, see Putney and McKay (1999) *BioEssays* 21:38-46).

[0034] Members of the capacitative calcium channel family include the calcium release-activated calcium current (CRAC) (Hoth and Penner (1992) *Nature* 355: 353-355), calcium release-activated nonselective cation current (CRANC) (Krause et al. (1996) *J. Biol. Chem.* 271: 32523-32528), and the transient receptor potential (TRP) proteins. There is no single electrophysiological profile characteristic of the family; rather, a wide array of single channel conductances, cation selectivity, and current properties have been observed for different specific channels. Further, in several instances it has been demonstrated that homo- or heteropolymerization of the channel molecule may occur, further changing the channel properties from that of the single molecule. In general, though, these channels function similarly, in that they are calcium ion-permeable cation channels which become activated upon stimulation of phospholipase C_β by a G protein-coupled receptor. Depletion of intracellular calcium stores activate these channels by a mechanism which is as yet undefined, but which has been demonstrated to involve a diffusible factor using studies in which calcium stores were artificially depleted (e.g., by the introduction of chelators into the cell, by activating phospholipase C_γ, or by inhibiting the those enzymes responsible for pumping calcium ions into the stores or those enzymes responsible for maintaining resting intracellular calcium ion concentrations) (Putney, J. W., (1986) *Cell Calcium* 7: 1-12; Putney, J. W. (1990) *Cell Calcium* 11:611-624).

[0035] The TRP channel family is one of the best characterized of the capacitative calcium channel group. These channels include transient receptor potential protein and homologues thereof (to date, seven homologs and splice variants have been identified in a variety of organisms), the vanilloid receptor subtype I (also known as the capsaicin receptor), stretch-inhibitable non-selective cation channel (SIC), olfactory, mechanosensitive channel, insulin-like growth factor I-regulated calcium channel, and vitamin D-responsive apical, epithelial calcium channel (ECaC) (see, e.g., Montell and Rubin (1989) *Neuron* 2:1313-1323; Caterina et al. (1997) *Nature* 389: 816-824; Suzuki et al. (1999) *J. Biol. Chem.* 274: 6330-6335; Kiselyov et al. (1998) *Nature* 396: 478-482; and Hoenderop et al. (1999) *J. Biol. Chem.* 274: 8375-8378). Each of these molecules is 700 or more amino acids (TRP and TRP homologs have 1300 or more amino acid residues), and shares certain conserved structural features. Predominant among these structural features are six transmembrane domains, with an additional hydrophobic loop present between the fifth and sixth transmembrane domains. It is believed that this loop is integral to the activity of the pore of the channel formed upon membrane insertion (Hardie and

Minke (1993) *Trends Neurosci* 16: 371-376). TRP channel proteins also include one or more ankyrin domains and frequently display a proline-rich region at the N-terminus. Although found in disparate tissues and organisms, members of the TRP channel protein family all serve to transduce signals by means of calcium entry into cells, particularly pain (see, e.g., McClesky and Gold (1999) *Annu. Rev. Physiol.* 61: 835-856), light (Hardie and Minke, supra), or olfactory signals (Colbert et al. (1997) *J. Neurosci* 17(21): 8259-8269). Thus, this family of molecules may play important roles in sensory signal transduction in general.

[0036] Calcium signaling may play a role in liver disease. Ca^{2+} influx has been shown to be essential for the contractile phenotype of activated stellate cells, being the phenotype considered responsible for the high portal hypertension associated with hepatic fibrosis. Hepatic stellate cells, a scarce liver cell type, have been proposed as the main effector of the fibrotic process. Once stimulated, stellate cells acquire the activated phenotype, proliferate, and become fibrogenic. Activated stellate cells contribute to the build-up of extracellular matrix (ECM) via overproduction of ECM components (e.g., collagen), and inhibition of their breakdown. The stimuli for stellate cell activation are not yet clear, although inflammatory cells (e.g., T-lymphocytes) and their mediators (e.g., growth factors, cytokines, and chemokines) interacting with their specific receptors (e.g., GPCRs), have all been postulated to play a role. In addition, PDGF-mediated stellate cell proliferation (a key phenotype of activated stellate cells) depends on Ca^{2+} influx.

[0037] As the TLCC molecules of the present invention may modulate calcium channel mediated activities, they may be useful for developing novel diagnostic and therapeutic agents for calcium channel associated disorders.

[0038] As used herein, a "calcium channel associated disorder" includes a disorder, disease or condition which is characterized by a misregulation of calcium channel mediated activity. Calcium channel associated disorders include cardiovascular disease and hepatic disorders. A cardiovascular disease or disorder also includes an endothelial cell and/or smooth muscle cell disorder.

[0039] Calcium channel disorders may also include CNS disorders and pain disorders. Pain disorders include those that affect pain signaling mechanisms.

[0040] Calcium channel disorders also include cellular proliferation, growth, differentiation, or migration disorders. The TLCC molecules of the present invention are involved in signal transduction mechanisms, which are known to be involved in cellular growth, differentiation, and migration processes. Thus, the TLCC molecules may modulate cellular growth, differentiation, or migration, and may play a role in disorders characterized by aberrantly regulated growth, differentiation, or migration.

[0041] As used herein, a "calcium channel mediated activity" includes an activity which involves a calcium channel, e.g., a calcium channel in a neuronal cell, a muscular cell, a vascular cell, or a liver cell, associated with receiving, conducting, and transmitting signals, in, for example, the nervous system. Calcium channel mediated activities include release of neurotransmitters or second messenger molecules (e.g., dopamine or norepinephrine), from cells, e.g., neuronal cells; modulation of resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation; participation in signal transduction pathways, and modulation of processes such as integration of sub-threshold synaptic responses and the conductance of back-propagating action potentials in, for example, neuronal cells (e.g., changes in those action potentials resulting in a morphological or differentiative response in the cell).

[0042] The term "family" when referring to the protein and nucleic acid molecules of the invention is intended to mean two or more proteins or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein. For example, the family of TLCC proteins comprises at least one "transmembrane domain" and preferably six transmembrane domains. As used herein, the term "transmembrane domain" includes an amino acid sequence of about 20-45 amino acid residues in length which spans the plasma membrane. More preferably, a transmembrane domain includes about at least 20, 25, 30, 35, 40, or 45 amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an alpha-helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, e.g., leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains are described in, for example, Zagotta W. N. et al, (1996) *Annual Rev. Neurosci.* 19: 235-263, the contents of which are incorporated herein by reference. Amino acid residues 599-619, 690-712, 784-803, 811-831, 845-862, and 933-957 of the TLCC protein (SEQ ID NO:2) comprise transmembrane domains. Accordingly, TLCC proteins having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with a transmembrane domain of human TLCC are within the scope of the invention.

[0043] In another embodiment, a TLCC molecule of the present invention is identified based on the presence of at least one pore domain between the fifth and sixth transmembrane domains. As used herein, the term "pore domain" includes an overall hydrophobic amino acid sequence which is located between two transmembrane domains of a calcium channel protein, preferably transmembrane domains 5 and 6, and which is believed to be a major determinant of ion selectivity and channel activity in calcium channels. Pore domains are described, for example in Vannier et al. (1998) *J. Biol. Chem.* 273: 8675-8679 and Phillips, A. M. et al. (1992) *Neuron* 8, 631-642, the contents of which are incorporated herein by reference. Amino acid residues 880-900 of the TLCC protein (SEQ ID NO:2) comprise a pore domain.

[0044] In another embodiment, a TLCC molecule of the present invention is identified based on the presence of at least one N-glycosylation site. As used herein, the term "N-glycosylation site" includes an amino acid sequence of about 4 amino acid residues in length which serves as a glycosylation site. More preferably, an N-glycosylation site has the consensus sequence Asn-Xaa-Ser/Thr-Xaa (where Xaa may be any amino acid except proline) (SEQ ID NO:4). N-glycosylation sites are described in, for example, Prosite PDOC00001. Amino acid residues 143-146, 205-208, and 907-910 of the TLCC protein (SEQ ID NO:2) comprise N-glycosylation sites. Accordingly, TLCC proteins having at least one N-glycosylation site are within the scope of the invention.

[0045] In another embodiment, a TLCC molecule of the present invention is identified based on the presence of a "transmembrane calcium channel domain" in the protein or corresponding nucleic acid molecule. As used herein, the term "transmembrane calcium channel domain" includes a

protein domain having an amino acid sequence of about 40-100 amino acid residues and having a bit score for the alignment of the sequence to the transmembrane calcium channel domain of at about 50-100. Preferably, a transmembrane calcium channel domain includes at least about 60-80, or more preferably about 63 amino acid residues, and has a bit score for the alignment of the sequence to the transmembrane calcium channel domain of at least 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70 or higher. The transmembrane calcium channel domain has been assigned ProDom entry 2328. To identify the presence of a transmembrane calcium channel domain in a TLCC protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein is searched against a database of known protein domains (e.g., the ProDom database) using the default parameters. A search was performed against the ProDom database resulting in the identification of a transmembrane calcium channel domain in the amino acid sequence of human TLCC (SEQ ID NO: 2) at about residues 783-845 of SEQ ID NO: 2.

[0046] Isolated proteins of the present invention, preferably TLCC proteins, have an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:2 or are encoded by a nucleotide sequence sufficiently identical to SEQ ID NO:1 or 3. As used herein, the term "sufficiently identical" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (e.g., an amino acid residue which has a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences share common structural domains or motifs and/or a common functional activity. For example, amino acid or nucleotide sequences which share common structural domains have at least 30%, 40%, or 50% homology, preferably 60% homology, more preferably 70%-80%, and even more preferably 90-95% homology across the amino acid sequences of the domains and contain at least one and preferably two structural domains or motifs, are defined herein as sufficiently identical. Furthermore, amino acid or nucleotide sequences which share at least 30%, 40%, or 50%, preferably 60%, more preferably, 70-80%, or 90-95% homology and share a common functional activity are defined herein as sufficiently identical.

[0047] As used interchangeably herein, an "TLCC activity", "biological activity of TLCC" or "functional activity of TLCC", refers to an activity exerted by a TLCC protein, polypeptide or nucleic acid molecule on a TLCC responsive cell or tissue, or on a TLCC protein substrate, as determined in vivo, or in vitro, according to standard techniques. In one embodiment, a TLCC activity is a direct activity, such as an association with a TLCC-target molecule. As used herein, a "target molecule" or "binding partner" is a molecule with which a TLCC protein binds or interacts in nature, such that TLCC-mediated function is achieved. A TLCC target molecule can be a non-TLCC molecule or a TLCC protein or polypeptide of the present invention. In an exemplary embodiment, a TLCC target molecule is a TLCC ligand, e.g., a calcium channel ligand. Alternatively, a TLCC activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the TLCC protein with a TLCC ligand. The biological activities of TLCC are described herein. For example, the TLCC proteins of the present invention can have one or more of the following activities: (1) modulate membrane excitability, (2) influence the resting potential of membranes, (3) modulate wave forms and frequencies of action potentials, (4) modulate thresholds of excitation, (5) modulate neurite outgrowth and synaptogenesis, (6) modulate signal transduction, (7) participate in nociception, (8) modulate hepatic disorders, (9) modulate angiogenesis, (10) modulate endothelial cell proliferation, and (11) modulate vascular tone.

[0048] Accordingly, another embodiment of the invention features isolated TLCC proteins and polypeptides having a TLCC activity. Preferred proteins are TLCC proteins having at least one transmembrane domain, and, preferably, a TLCC activity. Other preferred proteins are TLCC proteins having an N-glycosylation site and, preferably, a TLCC activity. Yet other preferred proteins are TLCC proteins having at least one transmembrane calcium channel domain and, preferably, a TLCC activity. Yet other preferred proteins are TLCC proteins having at least one transmembrane domain, at least one N-glycosylation site, and a transmembrane calcium channel domain and, preferably, a TLCC activity.

[0049] Additional preferred proteins have at least one transmembrane domain, and one or more of the following domains: at least one N-glycosylation site, and a transmembrane calcium channel domain, and are, preferably, encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3.

Tissue Distribution of TLCC mRNA

[0050] This example describes the tissue distribution of TLCC mRNA, as was qualitatively determined by Polymerase Chain Reaction (PCR), and quantitatively measured using the Taqman™ procedure.

[0051] Using PCR techniques, the human TLCC gene was determined to be predominantly expressed in osteoblasts, with some expression also seen in brain, adipose tissue, breast, colon, all fetal tissues, liver, pituitary, melanocyte, prostate, cervix, muscle, small intestine, megakaryocytes, and aorta, as well as in lymphoma and colon to liver metastases.

[0052] Using the Taqman™ procedure, it was determined that TLCC mRNA was expressed at low levels in normal human heart, kidney, lung, and liver. A very marked upregulation was detected in passaged human stellate cells, as well as in human fibrotic livers, although expression was low in quiescent stellate cells. TLCC mRNA was upregulated in human dermal and lung fibroblasts cultured in the presence of TGF-β.

[0053] It was also determined that the rat orthologue of TLCC was highly increased in all bile duct ligation-induced fibrotic livers tested as compared to control animals. Upregulation was detected in all carbon tetrachloride-induced fibrotic livers as compared to controls. However, there was no significant regulation in the serum-induced fibrotic livers as compared to controls, and no regulation in the cultured rat stellate cells. These data reveal that TLCC is highly regulated in activated stellate cells and in fibrotic livers, being expressed only at low levels in other organs and cell types. These observations suggest that TLCC may play an important role in Ca^{2+}-dependent phenomena (e.g., hepatic cell contractility and proliferation). The functional linkage of TRP channels to inositol triphosphate further suggests that TLCC might be related to key signaling events during stellate cell activation.

Expression of TLCC in Blood Vessels

[0054] Reverse Transcriptase PCR (RT-PCR) was performed using the Taqman procedure to detect the presence of

RNA transcripts corresponding to human TLCC in mRNA prepared from isolated human vessels or cells cultured from the endothelial vasculature. Significant TLCC expression was detected in vascular smooth muscle cells cultured from human aorta as well as in endothelial cells cultured from lung microvasculature or umbilical vein. Expression of TLCC was downregulated when cultured umbilical vein endothelial cells were treated with human recombinant IL-1β for six hours. Expression of TLCC in several isolated human vessels exceeded the expression level of TLCC in human adipose tissue which was included as a control.

Expression of TLCC in Endothelial Cells During Laminar Shear Stress

[0055] Human umbilical vein endothelial cells (HUVECs) were cultured in vitro under standard conditions, described in, for example, U.S. Pat. No. 5,882,925. Experimental cultures were then exposed to laminar shear stress (LSS) conditions.

[0056] Cultured HUVEC monolayers were exposed to laminar sheer stress by culturing the cells in a specialized apparatus containing liquid culture medium. Static cultures grown in the same medium served as controls. The in vitro LSS treatment at 10 dyns/cm^2 was performed for 24 hours and was designed to simulate the shear stress generated by blood flow in a straight, healthy artery.

[0057] The effect of LSS on TLCC expression in endothelial cells was assessed from total RNA prepared from the cells and used to probe clones arrayed on nylon filters. A TLCC clone showed a higher signal when probed with two of the three LSS samples when compared to their static controls, indicating that expression of TLCC is upregulated by laminar shear stress.

Human 15603

[0058] The present invention relates to a human TRP6, 15603. *Drosophila* transient receptor potential (TRP) proteins and some mammalian homologues (TRPC proteins) are thought to mediate capacitative Ca^{+2} entry (Hofmann et al., *Nature* 397:259-263 (1999)). Seven mammalian homologous genes (TRPC 1 to TRPC 7) have been cloned and characterized. TRPC6, together with TRPC 3 have been identified as the first members of a new functional family of second-messenger-operated cation channels, which are activated by diacylglycerol independently of protein kinase C (Hofmann et al., supra). It has recently been demonstrated that TRPC6 is likely to be the essential component of the α1-adrenoceptor-activated nonselective cation channel (α1-AR-NSCC), which may serve as a store-depletion-independent Ca^{2+} entry pathway during increased sympathetic activity. The α1-androgen receptor (α1-AR) is expressed widely in the vascular system. α1-AR stimulation leads to activation of G protein-coupled phospholipase Cβ, which catalyzes formation from phosphoinositide of 1,4,5-triphosphate (IP3) and diacylglycerol, which leads to the release of stored Ca^{2+} and sustained Ca^{2+} entry.

[0059] The human 15603 sequence (SEQ ID NO:5), which is approximately 2796 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 2793 nucleotides (nucleotides 1-2793 of both SEQ ID NO:5 and SEQ ID NO:7) not including the terminal codon. The coding sequence encodes a 931 amino acid protein (SEQ ID NO:6).

[0060] The human 15603 protein of SEQ ID NO:6 includes an amino-terminal hydrophobic amino acid sequence, consistent with a signal sequence, of about 26 amino acids (from amino acid 1 to about amino acid 26 of SEQ ID NO:6, PSORT, Nakai and Kanehisa (1992) *Genomics* 14:897-911), which upon cleavage results in the production of a mature protein form. This mature protein form is approximately 905 amino acid residues in length (from about amino acid 27 to amino acid 931 of SEQ ID NO:6).

[0061] Human 15603 contains the following regions or other structural features (for general information regarding PFAM identifiers, PS prefix and PF prefix domain identification numbers, refer to Sonnhammer et al. (1997) *Protein* 28:405-420: an ank domain (PFAM Accession Number PF0023) located at about amino acid residues 97 to 131 of SEQ ID NO:6; an ank domain (PFAM Accession Number PF0023) located at about amino acid residues 132 to 163 of SEQ ID NO:6; an ank domain (PFAM Accession Number PF0023) located at about amino acid residues 218 to 250 of SEQ ID NO:6; an ion transport protein domain (PFAM Accession Number PF00520) located at about amino acid residues 493 to 727 of SEQ ID NO:6; a Receptor Channel Potential Transient Repeat Variant Calcium Entry Ionic domain (ProDom No. PD140156) located at about amino acid residues 1 to 39 of SEQ ID NO:6; a Channel Receptor Transient Potential Calcium Repeat Variant Entry Ion domain (ProDom No. PD323618) located at about amino acid residues 40 to 71 of SEQ ID NO:6; a Channel Receptor Calcium Potential Capacitative Entry Trp-Related Transient Repeat Variant domain (ProDom No. PD186301) located at about amino acid residues 93 to 133 of SEQ ID NO:6; a Repeat Channel Ionic Receptor Trp Vision Transient Potential Ank domain (ProDom No. PD140149) located at about amino acid residues 95 to 154 of SEQ ID NO:6; a Repeat Ankyrin Channel Gene ORF Family Receptor Ankyrin-Like Factor (ProDom No. PD007334) located at about amino acid residues 102 to 166 of SEQ ID NO:6; a Ankyrin Repeat Kinase Domain UNC-44 Ankyrin-Related Alternative Glycoprotein EGF-like domain (ProDom No. PD000041) located at about amino acid residues 105 to 188 of SEQ ID NO:6; a Channel Receptor Transient Repeat Calcium Potential Ion Ank Transport Transmembrane domain (ProDom No. PD004194) located at about amino acid residues 123 to 347 of SEQ ID NO:6; a Receptor Channel Potential Transient NOMPC TRP2Y71A12B.4Y71A12B.E 2-Beta 2-Alpha domain (ProDom No. PD296552) located at about amino acid residues 299 to 509 of SEQ ID NO:6; a Transmembrane Fis Receptor MTR1 domain (ProDom No. PD039592) located at about amino acid residues 308 to 916 of SEQ ID NO:6; a Channel Receptor Calcium Repeat Transient Potential Ion Transmembrane Ionic Transport domain (ProDom No. PD328255) located at about amino acid residues 362 to 509 of SEQ ID NO:6; a Channel Receptor Calcium Transient Potential Repeat Vanilloid Transmembrane Ion Transport domain (ProDom No. PD003230) located at about amino acid residues 409 to 748 of SEQ ID NO:6; a Channel Receptor Calcium Potential Repeat Transient Ion Ionic Variant Ank domain (ProDom No. PD238062) located at about amino acid residues 510 to 597 of SEQ ID NO:6; a Channel Receptor Calcium Repeat Potential Transient Capacitative Entry Ion Transport domain (ProDom No. PD342728) located at about amino acid residues 744 to 895 of SEQ ID NO:6; a Channel Receptor Transient Potential Repeat Calcium Transport Transmembrane Ank Ionic domain (ProDom No. PD004174)

located at about amino acid residues 749 to 900 of SEQ ID NO:6; a Channel Repeat Calcium Ionic Entry Receptor Transient Ank Transport domain (ProDom No. PD266294) located at about amino acid residues 786 to 854 of SEQ ID NO:6; a Gelsolin-Related domain (ProDom No. PD202783) located at about amino acid residues 792 to 931 of SEQ ID NO:6; a Coiled Coil Myosin Repeat Chain Heavy Filament Heptad Pattern Muscle domain (ProDom No. PD000002) located at about amino acid residues 796 to 928 of SEQ ID NO:6; a Channel Receptor Transient Potential Calcium Repeat Variant Entry Ion (ProDom No. PD137340) located at about amino acid residues 901 to 931 of SEQ ID NO:6; a transmembrane domain (predicted by MEMSAT, Jones et al. (1994) *Biochemistry* 33:3038-3049) at about amino acids 407 to 426 of SEQ ID NO:6; a transmembrane domain (predicted by MEMSAT, Jones et al. (1994) *Biochemistry* 33:3038-3049) at about amino acids 443 to 459 of SEQ ID NO:6; a transmembrane domain (predicted by MEMSAT, Jones et al. (1994) *Biochemistry* 33:3038-3049) at about amino acids 490 to 507 of SEQ ID NO:6; a transmembrane domain (predicted by MEMSAT, Jones et al. (1994) *Biochemistry* 33:3038-3049) at about amino acids 598 to 614 of SEQ ID NO:6; a transmembrane domain (predicted by MEMSAT, Jones et al. (1994) *Biochemistry* 33:3038-3049) at about amino acids 637 to 653 of SEQ ID NO:6; a transmembrane domain (predicted by MEMSAT, Jones et al. (1994) *Biochemistry* 33:3038-3049) at about amino acids 703 to 727 of SEQ ID NO:6; two coiled coil segments located at about amino acids 297 to 327 (ALELSNELAVLANIEKEFKNDYKKLSMQCKD; SEQ ID NO:88) and 878 to 923 (EVNEGELKEIKQDISSLRYELLEEKSQNTEDLAELIRELGEKLSME; SEQ ID NO:89) of SEQ ID NO:6; a leucine zipper pattern (Prosite PS00029) located at about amino acids 766 to 787 (LVPSPKSLFYLLLKLKKWISEL; SEQ ID NO:90) of SEQ ID NO:6; nine protein kinase C phosphorylation sites (Prosite PS00005) located at about amino acids 14 to 16 (SPR), 89 to 91 (SDR), 563 to 565 (TLK), 630 to 632 (TVK), 674 to 676 (SFK), 769 to 771 (SPK), 836 to 838 (SIR), 893 to 895 (SLR), and 929 to 931 (TNR) of SEQ ID NO:6; sixteen casein kinase II phosphorylation sites (Prosite PS00006) located at about amino acids 96 to 99 (SIEE; SEQ ID NO:91), 197 to 200 (SQSE; SEQ ID NO:92), 216 to 219 (SSHD; SEQ ID NO:93), 289 to 292 (SSED; SEQ ID NO:94), 296 to 299 (TALE; SEQ ID NO:95), 475 to 478 (TSTD; SEQ ID NO:96), 492 to 495 (SWME; SEQ ID NO:97), 556 to 559 (SIID; SEQ ID NO:98), 563 to 566 (TLKD; SEQ ID NO:99), 571 to 574 (TLGD; SEQ ID NO:100), 630 to 633 (TVKD; SEQ ID NO:101), 669 to 672 (TTVE; SEQ ID NO:102), 730 to 733 (SFQE; SEQ ID NO:103), 752 to 755 (SYFE; SEQ ID NO:104), 815 to 818 (SHED; SEQ ID NO:105), and 839 to 842 (SSED; SEQ ID NO:106) of SEQ ID NO:6; two cAMP/cGMP-dependent protein kinase phosphorylation sites (Prosite PS00004) located at about amino acids 67 to 70 (RRQT; SEQ ID NO:107) and 319 to 322 (KKLS; SEQ ID NO:108) of SEQ ID NO:6; three tyrosine phosphorylation sites (Prosite PS00007) located at about amino acids 24 to 31 (RRNESQDY; SEQ ID NO:109), 101 to 108 (RFLDAAEY; SEQ ID NO:110), and 698 to 705 (KFIENIGY; SEQ ID NO:111) of SEQ ID NO:6; two amidation sites (Prosite PS00009) located at about amino acids 75 to 78 (KGRR; SEQ ID NO:112) and 188 to 191 (EGKR; SEQ ID NO:113) of SEQ ID NO:6; nine N-glycosylation sites (Prosite PS00001) located at about amino acids 26 to 29 (NESQ; SEQ ID NO:114), 157 to 160 (NLSR; SEQ ID NO:115), 362 to 365 (NLSR; SEQ ID NO:116), 394 to 397 (NLSG; SEQ ID NO:117), 473 to 476 (NETS; SEQ ID NO:118), 561 to 564 (NDTL; SEQ ID NO:119), 617 to 620 (NESF; SEQ ID NO:120) 712 to 715 (NVTM; SEQ ID NO:121) and 728 to 731 (NSSF; SEQ ID NO:122) of SEQ ID NO:6; and five N-myristoylation sites (Prosite PS00008) located at about amino acids 17 to 22 (GAAGAA; SEQ ID NO:123), 213 to 218 (GTRSSH; SEQ ID NO:124), 504 to 509 (GMIWAE; SEQ ID NO:125), 661 to 666 (GAKQNE; SEQ ID NO:126), and 811 to 816 (GILGSH; SEQ ID NO:127) of SEQ ID NO:6.

[0062] A hydropathy plot of human 15603 was performed. Polypeptides of the invention include fragments which include: all or part of a hydrophobic sequence, e.g., the sequence from about amino acid 162 to 172, from about amino acid 215 to 225, from about amino acid 325 to 335, from about amino acid 401 to 431, from about amino acid 441 to 461, from about amino acid 486 to 506, from about amino acid 521 to 551, from about amino acid 591 to 616, from about amino acid 631 to 656, from about amino acid 665 to 675, from about amino acid 701 to 731, and from about amino acid 770 to 780 of SEQ ID NO:6; all or part of a hydrophilic sequence, e.g., the sequence from about amino acid 20 to 32, from about amino acid 55 to 85, from about amino acid 181 to 221, from about amino acid 241 to 281, from about amino acid 306 to 321, from about amino acid 331 to 340, from about amino acid 461 to 486, from about amino acid 780 to 805, from about amino acid 810 to 841, from about amino acid 841 to 851, from about amino acid 865 to 890, from about amino acid 895 to 905, and from about amino acid 915 to 931 of SEQ ID NO:6; a sequence which includes a Cys, or a glycosylation site.

[0063] The 15603 protein contains a significant number of structural characteristics in common with members of the ion channel family and the ank repeat family. As used herein, the term "ion channel" includes a protein or polypeptide which is capable of regulating the flow of ions such as calcium cations through a channel. The channel may regulate the flow of a particular cation or anion or may be less discriminating and allow multiple types of cations or anions to flow through it. The flow of ions may be regulated by the presence or absence of a bound ligand or may be regulated by the phosphorylation state of the channel or another protein associated with the channel. The ion channel protein may undergo a conformational change based on the binding of a ligand, the phosphorylation of a particular residue, or the binding of another protein or biomolecule.

[0064] Members of an ion channel family of proteins are characterized by transmembrane domains, cytoplasmic domains, extracellular domains, and may be homodimers, homotrimers, homomultimers, heterodimeric etc. The pores of ion channels are typically formed by multiple transmembrane proteins encoded by the same or different genes. Typically, a hydrophilic transmembrane channel is formed that allows passage of ions from one side of the plasma membrane to the other. Clusters of charged amino acids at the mouth of the channel may increase the selectivity for a particular type of ion, e.g., clusters of negatively charged amino acid residues at the mouth of the channel may exclude negative ions from cation channels.

[0065] All ion channel family proteins form water-filled pores across membranes. Ion channel proteins are located in the plasma membrane of animal and plant cells and are further characterized by small highly selective pores that participate in ion transport. Typically, more than 10^6 ions can pass

through such a channel each second. Many ion channels allow specific ions, such as Na^+, K^+, Ca^{2+}, or Cl^-, to diffuse down their electrochemical gradients across the lipid bilayer. The ion channel proteins show ion selectivity, permitting some ions to pass but not others. Another feature of ion channel proteins is that they are not continuously open, in contrast to simple aqueous pores. Instead, they have "gates," which open briefly and then close again. This opening and closing is usually in response to a specific perturbation of the membrane, such as a change in voltage across the membrane (voltage-gated channels), mechanical stimulation (mechanically-gated channels), or the binding of a signaling molecule (ligand-gated channels). In the case of ligand-gated channels, the signaling ligand can be either an extracellular mediator, such as a neurotransmitter (transmitter-gated channels), an intracellular mediator, such as an ion (ion-gated channels), a nucleotide (nucleotide-gated channels), or a GTP-binding regulatory protein (G-protein-gated channels).

[0066] Ion channels are responsible for the electrical excitability of nerve and muscle cells and mediate most forms of electrical signaling in the nervous system. A single nerve cell typically contains more than five kinds of ion channels. However, these ion channels are not restricted to electrically excitable cells. Ion channels are present in all animal cells and are found in plant cells and microorganisms. For example, ion channel proteins propagate the leaf-closing response of the mimosa plant and allow the single-celled paramecium to reverse direction after collision. 15603, also known as TRPC6, is a non-selective cation channel that is activated by diacylglycerol in a membrane-delimited fashion, independently of protein kinase C

[0067] A 15603 polypeptide can include a "ion channel domain" or regions homologous with a "ion channel domain". A 15603 polypeptide can further include a "ank domain" or regions homologous with a "ank domain".

[0068] As used herein, the term "ion transport protein domain" includes an amino acid sequence of about 200 to 250 amino acid residues in length and having a bit score for the alignment of the sequence to the ion transport protein domain (HMM) of at least 75. Preferably, an ion transport domain mediates the flow of ions through a membrane. Preferably, an ion transport protein domain includes at least about 200 to 300 amino acids, more preferably about 200 to 250 amino acid residues, or about 210 to 240 amino acids and has a bit score for the alignment of the sequence to the ion transport protein domain (HMM) of at least 50, 75, 80, or greater. The ion transport domain can include transmembrane domains. The ion transport protein domain (HMM) has been assigned the PFAM Accession Number PF00520. The ion transport protein domain (amino acids 493 to 727 of SEQ ID NO:6) of human 15603 aligns with the Pfam ion transport protein consensus amino acid sequence (SEQ ID NO:11) derived from a hidden Markov model.

[0069] As used herein, the term "ank repeat domain" includes an amino acid sequence of about 30 to 35 amino acid residues in length and having a bit score for the alignment of the sequence to the ank repeat domain (HMM) of at least 15. Preferably, an ank repeat domain includes at least about 20 to 40 amino acids, more preferably about 25 to 35 amino acid residues, or about 30 to 35 amino acids and has a bit score for the alignment of the sequence to the ank repeat domain (HMM) of at least 10, 15, 20, or greater. The ank repeat domain (HMM) has been assigned the PFAM Accession Number PF0023. Additionally, the ank repeat domain (HMM) has been assigned the SMART identifier ANK_2a. The ank repeat domains (amino acids 97 to 131, 132 to 163, and 218 to 250 of SEQ ID NO:6) of human 15603 align with the Pfam ank repeat domain consensus amino acid sequences (SEQ ID NO:8-10) derived from a hidden Markov model. The ank repeat domains (amino acids 97 to 126, 132 to 160, and 218 to 247 of SEQ ID NO:6) of human 15603 align with the SMART ANK_2a domain consensus amino acid sequences (SEQ ID NO:8-10) derived from a hidden Markov model.

[0070] In a preferred embodiment, a 15603 polypeptide or protein has an "ion transport protein domain" or a region which includes at least about 200 to 300 more preferably about 200 to 250 or 210 to 240 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "ion transport protein domain," e.g., the ion transport protein domain of human 15603 (e.g., residues 493 to 727 of SEQ ID NO:6).

[0071] In a preferred embodiment, a 15603 polypeptide or protein has an "ank repeat domain" or a region which includes at least about 25 to 40 more preferably about 25 to 35 or 30 to 35 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with an "ank repeat domain," e.g., an ank repeat domain of human 15603 (e.g., residues 97 to 131, 132 to 163, or 218 to 250 of SEQ ID NO:6).

[0072] To identify the presence of a "ion transport protein" domain or an "ank repeat" domain in a 15603 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against the Pfam database of HMMs (e.g., the Pfam database, release 2.1) using the default parameters. For example, the hmmsf program, which is available as part of the HMMER package of search programs, is a family specific default program for MILPAT0063 and a score of 15 is the default threshold score for determining a hit. Alternatively, the threshold score for determining a hit can be lowered (e.g., to 8 bits). A description of the Pfam database can be found in Sonhammer et al. (1997) *Proteins* 28:405-420 and a detailed description of HMMs can be found, for example, in Gribskov et al. (1990) *Meth. Enzymol.* 183:146-159; Gribskov et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:4355-4358; Krogh et al. (1994) *J. Mol. Biol.* 235:1501-1531; and Stultz et al. (1993) *Protein Sci.* 2:305-314, the contents of which are incorporated herein by reference. A search was performed against the HMM database resulting in the identification of an "ion transport protein" domain in the amino acid sequence of human 15603 at about residues 493 to 727 of SEQ ID NO:6. A search was performed against the HMM database resulting in the identification of three "ank repeat" domains in the amino acid sequence of human 15603 at about residues 97 to 131, 132 to 163, and 218 to 250 of SEQ ID NO:6.

[0073] An additional method to identify the presence of an "ank repeat" domain in a 15603 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against a SMART database (Simple Modular Architecture Research Tool) of HMMs as described in Schultz et al. (1998), *Proc. Natl. Acad. Sci. USA* 95:5857 and Schultz et al. (2000) *Nucl. Acids Res* 28:231. The database contains domains identified by profiling with the hidden Markov models of the HMMer2 search program (Durbin et al. (1998) *Biological sequence analysis: probabilistic models*

of proteins and nucleic acids. Cambridge University Press). The database also is extensively annotated and monitored by experts to enhance accuracy. A search was performed against the HMM database resulting in the identification of three "ank 2a" domains in the amino acid sequence of human 15603 at about residues 97 to 126, 132 to 160, and 218 to 247 of SEQ ID NO:6.

[0074] For further identification of domains in a 15603 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against a database of domains, e.g., the ProDom database (Corpet et al. (1999), *Nucl. Acids Res.* 27:263-267). The ProDom protein domain database consists of an automatic compilation of homologous domains. Current versions of ProDom are built using recursive PSI-BLAST searches (Altschul et al. (1997) *Nucleic Acids Res.* 25:3389-3402; Gouzy et al. (1999) *Computers and Chemistry* 23:333-340) of the SWISS-PROT 38 and TREMBL protein databases. The database automatically generates a consensus sequence for each domain. A BLAST search was performed against the HMM database resulting in the identification of a "channel receptor transient repeat calcium potential ion ank transport transmembrane" domain in the amino acid sequence of human 15603 at about residues 123 to 347 of SEQ ID NO:6. A BLAST alignment of the human 15603 ion channel domain with a consensus amino acid sequence of a domain derived from the ProDomain database ("Channel Receptor Transient Repeat Calcium Potential Ion Ank Transport Transmembrane;" No. PD004194; ProDomain Release 2001.1) showed the amino acid residues 52 to 278 of the 278 amino acid PD004194 consensus sequence (SEQ ID NO:12) aligned with the ion channel domain of human 15603, amino acid residues 123 to 347 of SEQ ID NO:6.

[0075] A 15603 molecule can further include a transmembrane domain or an ank repeat domain.

[0076] A 15603 polypeptide can include at least one, two, three, four, five, preferably six "transmembrane domains" or regions homologous with "transmembrane domains". As used herein, the term "transmembrane domain" includes an amino acid sequence of about 10 to 40 amino acid residues in length and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, e.g., at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, e.g., leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains typically have alpha-helical structures and are described in, for example, Zagotta et al., (1996) *Annual Rev. Neurosci.* 19:235-263, the contents of which are incorporated herein by reference. The transmembrane domains of human 15603 are located at about residues 407 to 426, 443 to 459, 490 to 507, 598 to 614, 637 to 653, and 703 to 727 of SEQ ID NO:6.

[0077] In a preferred embodiment, a 15603 polypeptide or protein has at least one, two, three, four, five, preferably six "transmembrane domain" or a region which includes at least about 12 to 35 more preferably about 14 to 30 or 15 to 25 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "transmembrane domain," e.g., the transmembrane domains of human 15603 (e.g., residues 407 to 426, 443 to 459, 490 to 507, 598 to 614, 637 to 653, or 703 to 727 of SEQ ID NO:6). The transmembrane domain of human 15603 can be visualized in a hydropathy plot as regions of about 15 to 25 amino acids where the hydropathy trace is mostly above the horizontal line.

[0078] To identify the presence of a "transmembrane" domain in a 15603 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be analyzed by a transmembrane prediction method that predicts the secondary structure and topology of integral membrane proteins based on the recognition of topological models (MEMSAT, Jones et al., (1994) *Biochemistry* 33:3038-3049).

[0079] A 15603 polypeptide can include at least one, two, three, four, five, six, preferably seven "non-transmembrane regions." As used herein, the term "non-transmembrane region" includes an amino acid sequence not identified as a transmembrane domain. The non-transmembrane regions in 15603 are located at about amino acids 1 to 406, 427 to 442, 460 to 489, 508 to 597, 615 to 636, 654 to 702, and 727 to 931 of SEQ ID NO:6. Non-transmembrane domains can be cytoplasmic or extracellular.

[0080] The non-transmembrane regions of 15603 include at least one, two, three, preferably four cytoplasmic regions. When located at the N-terminus, the cytoplasmic region is referred to herein as the "N-terminal cytoplasmic domain." As used herein, an "N-terminal cytoplasmic domain" includes an amino acid sequence having about 1 to 500, preferably about 1 to 450, more preferably about 1 to 425, or even more preferably about 1 to 410 amino acid residues in length, is located inside of a cell or within the cytoplasm of a cell. The C-terminal amino acid residue of an "N-terminal cytoplasmic domain" is adjacent to an N-terminal amino acid residue of a transmembrane domain in a 15603 protein. For example, an N-terminal cytoplasmic domain is located at about amino acid residues 1 to 406 of SEQ ID NO:6.

[0081] In a preferred embodiment, a 15603 polypeptide or protein has an N-terminal cytoplasmic domain or a region which includes about 1 to 450, preferably about 1 to 425, and more preferably about 1 to 410 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with an "N-terminal cytoplasmic domain," e.g., the N-terminal cytoplasmic domain of human 15603 (e.g., residues 1 to 406 of SEQ ID NO:6).

[0082] In another embodiment, a 15603 cytoplasmic region includes at least one, preferably two cytoplasmic loops. As used herein, the term "loop" includes an amino acid sequence which is not included within a phospholipid membrane, having a length of at least about 4, preferably about 5 to 95, more preferably about 6 to 35 amino acid residues, and has an amino acid sequence that connects two transmembrane domains within a protein or polypeptide. Accordingly, the N-terminal amino acid of a loop is adjacent to a C-terminal amino acid of a transmembrane domain in a 15603 molecule, and the C-terminal amino acid of a loop is adjacent to an N-terminal amino acid of a transmembrane domain in a 15603 molecule. As used herein, a "cytoplasmic loop" includes a loop located inside of a cell or within the cytoplasm of a cell. For example, a "cytoplasmic loop" can be found at about amino acid residues 460 to 489 and 615 to 636 of SEQ ID NO:6.

[0083] In a preferred embodiment, a 15603 polypeptide or protein has a cytoplasmic loop or a region which includes at least about 4, preferably about 5 to 40, and more preferably about 6 to 35 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a cytoplasmic loop, e.g., a cytoplasmic loop of human 15603 (e.g., residues 460 to 489 and 615 to 636 of SEQ ID NO:6).

[0084] In another embodiment, a 15603 non-transmembrane region includes at least one, two, preferably three non-cytoplasmic loops. As used herein, a "non-cytoplasmic loop" includes a loop located outside of a cell or within an intracellular organelle. Non-cytoplasmic loops include extracellular domains (i.e., outside of the cell) and intracellular domains (i.e., within the cell). When referring to membrane-bound proteins found in intracellular organelles (e.g., mitochondria, endoplasmic reticulum, peroxisomes microsomes, vesicles, endosomes, and lysosomes), non-cytoplasmic loops include those domains of the protein that reside in the lumen of the organelle or the matrix or the intermembrane space. For example, a "non-cytoplasmic loop" can be found at about amino acid residues 427 to 442, 508 to 597, and 654 to 702 of SEQ ID NO:6.

[0085] In a preferred embodiment, a 15603 polypeptide or protein has at least one non-cytoplasmic loop or a region which includes at least about 4, preferably about 5 to 100, more preferably about 6 to 90 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "non-cytoplasmic loop," e.g., at least one non-cytoplasmic loop of human 15603 (e.g., residues 427 to 442, 508 to 597, and 654 to 702 of SEQ ID NO:6).

[0086] In another embodiment, a cytoplasmic region of a 15603 protein can include the C-terminus and can be a "C-terminal cytoplasmic domain," also referred to herein as a "C-terminal cytoplasmic tail." As used herein, a "C-terminal cytoplasmic domain" includes an amino acid sequence having a length of at least about 200, preferably about 150 to 250, more preferably about 175 to 225 amino acid residues, is located inside of a cell or within the cytoplasm of a cell. The N-terminal amino acid residue of a "C-terminal cytoplasmic domain" is adjacent to a C-terminal amino acid residue of a transmembrane domain in a 15603 protein. For example, a C-terminal cytoplasmic domain is located at about amino acid residues 728 to 931 of SEQ ID NO:6.

[0087] In a preferred embodiment, a 15603 polypeptide or protein has a C-terminal cytoplasmic domain or a region which includes at least about 200, preferably about 150 to 250, and more preferably about 175 to 225 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a C-terminal cytoplasmic domain, e.g., the C-terminal cytoplasmic domain of human 15603 (e.g., residues 728 to 931 of SEQ ID NO:6).

[0088] A human 15603 protein can further include a leucine zipper or coiled coil structure. Preferably a leucine zipper domain has at least about 60%, 70%, 80%, 90%, 95%, 99%, or 100% homology with the leucine zipper domain of human 15603 (e.g., residues 766 to 787 of SEQ ID NO:2). Preferably a coiled coil domain has at least about 60%, 70%, 80%, 90%, 95%, 99%, or 100% homology with the coiled coil domain of human 15603 (e.g., residues 296 to 907 of SEQ ID NO:6). A human 15603 protein can further include N-glycosylation sites, cAMP and cGMP-dependent protein kinase phosphorylation sites, protein kinase C phosphorylation sites, casein kinase II phosphorylation site, tyrosine kinase phosphorylation sites, N-myristoylation sites, and amidation sites.

[0089] A 15603 family member can include at least one ion channel protein domain; and at least one, two, preferably three ank repeat domains or transmembrane or non-transmembrane domains. A 15603 family member can include at least one leucine zipper structure or at least one coiled coil structure. Furthermore, a 15603 family member can include at least one, two, three, four, five, six, seven, eight, preferably nine protein kinase C phosphorylation sites (Prosite PS00005); at least one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, and preferably sixteen casein kinase II phosphorylation sites (Prosite PS00006); at least one, two, three, four, five, six, seven, eight, preferably nine N-glycosylation sites (Prosite PS00001); at least one, preferably two cAMP/cGMP protein kinase phosphorylation sites (Prosite PS00004); at least one, preferably two amidation sites (Prosite PS00009); and at least one, two, three, four, and preferably five N-myristoylation sites (Prosite PS00008).

[0090] As the 15603 polypeptides of the invention can modulate 15603-mediated activities, they can be useful for developing novel diagnostic and therapeutic agents for ion channel-associated or other 15603-associated disorders, as described below.

[0091] As used herein, a "ion channel-associated activity" includes an activity which involves the regulation of the flow of ions across a membrane. The flow of ions may be controlled by a second-messenger such as diacylglycerol. Members of the family can play a role in cardiovascular disease such as abnormal angiogenesis.

[0092] As used herein, a "15603 activity", "biological activity of 15603" or "functional activity of 15603", refers to an activity exerted by a 15603 protein, polypeptide or nucleic acid molecule on e.g., a 15603-responsive cell or on a 15603 substrate, e.g., a protein substrate, as determined in vivo or in vitro. In one embodiment, a 15603 activity is a direct activity, such as an association with a 15603 target molecule. A "target molecule" or "binding partner" is a molecule with which a 15603 protein binds or interacts in nature. In an exemplary embodiment, 15603 is a ion channel, e.g., a non-selective cation channel that is activated by diacylglycerol, and thus interacts in nature with a molecule such as a second messenger (e.g., diacylglycerol) to regulate the flow of cations through a membrane.

[0093] A 15603 activity can also be an indirect activity, e.g., a cellular signaling activity mediated by interaction of the 15603 protein with a 15603 receptor. Based on the above-described sequence structures and similarities to molecules of known function, the 15603 molecules of the present invention can have similar biological activities as ion channel family members. For example, the 15603 proteins of the present invention can have one or more of the following activities: (1) the ability to bind a second messenger; (2) the ability to bind diacylglycerol; (3) the ability to regulate the flow of cations through a membrane; (4) the ability to regulate angiogenesis; and (5) the ability to regulate intracellular calcium levels.

[0094] The 15603 molecules of the invention can modulate the activities of cells in tissues where they are expressed. For example, using TaqMan analysis, 15603 mRNA is expressed in moderate to high levels in hemangiomas, megakaryocytes, Wilm's tumor, vascular smooth muscle cells, proliferative endothelial cells, uterine, and normal brain cortex. Accordingly, the 15603 molecules of the invention can act as therapeutic or diagnostic agents for cardiovascular, angiogenic, or neoproliferative disorders.

[0095] Thus, the 15603 molecules can act as novel diagnostic targets and therapeutic agents for controlling one or more cardiovascular, angiogenic, neoproliferative, or other ion channel disorders. As used herein, "ion channel disorders" are diseases or disorders whose pathogenesis is caused by, is related to, or is associated with aberrant or deficient ion

channel protein function or expression. Examples of such disorders, e.g., ion channel-associated or other 15603-associated disorders, include but are not limited to, cellular proliferative and/or differentiative disorders, immune e.g., inflammatory, disorders, cardiovascular disorders, endothelial cell disorders, or renal disorders.

[0096] The 15603 molecules can be used to treat cardiovascular, immune system, or proliferative disorders in part because family members are found in the hemangiomas, blood vessels, and megakaryocytes.

[0097] The 15603 molecules of the invention can be used to monitor, treat and/or diagnose a variety of proliferative disorders. Such disorders include hematopoietic neoplastic disorders.

[0098] Aberrant expression and/or activity of 15603 molecules can mediate disorders associated with bone metabolism.

[0099] The 15603 nucleic acid and protein of the invention can be used to treat and/or diagnose a variety of immune, e.g., inflammatory, (e.g. respiratory inflammatory) disorders

[0100] Additionally, 15603 molecules can play an important role in the etiology of certain viral diseases, including but not limited to Hepatitis B, Hepatitis C and Herpes Simplex Virus (HSV). Modulators of 15603 activity could be used to control viral diseases. The modulators can be used in the treatment and/or diagnosis of viral infected tissue or virus-associated tissue fibrosis, especially liver and liver fibrosis. Also, 15603 modulators can be used in the treatment and/or diagnosis of virus-associated carcinoma, especially hepatocellular cancer.

[0101] Additionally, 15603 can play an important role in the regulation of metabolism or pain disorders.

Gene Expression Analysis of 15603

[0102] Human 15603 expression was measured by TaqMan® quantitative PCR (Perkin Elmer Applied Biosystems) in cDNA prepared from a variety of normal and diseased (e.g., cancerous) human tissues or cell lines.

[0103] The results indicate significant 15603 expression in hemangioma, vascular smooth muscle, megakaryocytes, Wilm's tumor; at medium levels in normal brain cortex, uterine, and proliferative endothelial cells; and at low levels in adipose tissue, normal human internal mammary artery, diseased human aorta, diseased human artery, diseased human vein, and normal human vein, normal heart, CHF heart, normal kidney, skeletal muscle, pancreas, primary osteoblasts, brain hypothalamus, normal breast, normal ovary, normal prostate, prostate tumor, salivary glands, normal colon, colon tumor, normal lung, lung tumor, lung COPD, colon IBD, normal spleen, normal tonsil, normal small intestine, skin decubitus, synovium, glioblastomas, fetal adrenal, fetal kidney, and fetal heart.

TABLE 2

Phase 1.5.1 Expression of 15603 w/beta 2

Tissue Type	Expression
Artery normal	0.6556
Aorta diseased	0.3871
Vein normal	0.1012
Coronary SMC	0
HUVEC	0
Hemangioma	28.1641

TABLE 2-continued

Phase 1.5.1 Expression of 15603 w/beta 2

Tissue Type	Expression
Heart normal	0.2302
Heart CHF	0.0759
Kidney	0.618
Skeletal Muscle	0.1668
Adipose normal	0.0932
Pancreas	0.172
Primary osteoblasts	0.1037
Osteoclasts (diff)	0
Skin normal	0
Spinal cord normal	0
Brain Cortex normal	2.2203
Brain-Hypothalamus normal	0.4603
Nerve	0
DRG (Dorsal Root Ganglion)	0
Breast normal	0.3587
Breast tumor	0
Ovary normal	0.7609
Ovary Tumor	0
Prostate Normal	0.0355
Prostate Tumor	0.1041
Salivary glands	0.0261
Colon normal	0.0164
Colon Tumor	0.0566
Lung normal	0.6053
Lung tumor	0.9175
Lung COPD	0.6978
Colon IBD	0.006
Liver normal	0
Liver fibrosis	0
Spleen normal	0.0303
Tonsil normal	0.0084
Lymph node normal	0
Small intestine normal	0.0894
Skin-Decubitus	0.3123
Synovium	0.0708
BM-MNC	0
Activated PBMC	0
Neutrophils	0
Megakaryocytes	17.3972
Erythroid	0

TABLE 3

15603 Expression in Vessel Panel 2.1

Tissue Type	Expression
Coronary SMC	0
Huvec NS	0
Huvec Shear/static pooled	0
H/Adipose/PIT 695	0.1151
H/Internal Mam Artery/Normal/AMC 263	0
H/Internal Mam Artery/Normal/AMC 347	0.3221
H/Internal Mam Artery/Normal/AMC350	0
H/Internal Mam Artery/Normal/AMC 352	0.2863
H/Artery/Normal/PIT 1180	0.0511
H/Artery/normal/AMC 150	0
H/Artery/normal/PIT 912	0.0765
H/Artery/normal/NDR 352	0.045
H/Aorta/Diseased/PIT 710	0
H/Aorta/Diseased/PIT 712	0
H/Aorta/Diseased/PIT 732	0.1909
H/Artery/Diseased/iliac/NDR 753	0
H/Artery/Diseased/Tibial/PIT 679	0.1238
H/Vein/Normal/PIT 1010	0.3255
H/Vein/Normal/NDR 239	0.0477
H/Vein/Normal/AMC 130	0.2581
H/Vein/Normal/AMC 131	0
H/Vein/Normal/AMC 137	0

TABLE 3-continued

15603 Expression in Vessel Panel 2.1	
Tissue Type	Expression
H/Vein/Normal/AMC 153	0.0661
H/Vein/Normal/AMC 176	0.0345
H/Vein/Normal/AMC 177	0
H/Vein/Normal/AMC 178	0.1002
H/Vein/Normal/AMC 182	0.0323
H/Vein/Normal/AMC 190	0
H/Vein/Normal/AMC 192	0
H/Vein/Normal/AMC 195	0
H/Vein/Normal/AMC 211	0
M/Aorta/Normal/PRI 286	0
M/Vein/Normal/PRI 328	0
M/Vein/Normal/PRI 230	0
M/Aorta/Diseased/CAR 1216	0
M/Aorta/Diseased/CAR 1237	0
M/Aorta/Diseased/CAR 1192	0
M/Aorta/Diseased/CAR 1196	0
M/Artery/Diseased/CAR 1174	6.3899
M/Artery/Diseased/CAR 1175	0

TABLE 4

Expression of 15063	
Tissue Type	Value
HUVECs Stat	0.
HUVECs LSS	0.
HUVECs Stat	0.
LSS	0.
HUVECs Prolif	0.
HUVECs Conf	0.
HUVECs -GF	0.
HUVECs IL-1	0.
HMVEC-Cardiac Prolif	0.
HMVEC-Cardiac Conf	0.
HMVEC-Cardiac Prolif	0.
HMVEC-Cardiac Conf	0.
HMVEC-Lung Prolif	0.
HMVEC-Lung Conf	0.
HMVEC-Lung -GF	0.
HMVEC-Lung Prolif	1.
HMVEC-Lung Conf	0.
Aortic C4h	0.
Aortic TNF	0.
Aortic C14	0.
Aortic TNF	0.
293	0.
HUVEC Stat	0.
HUVEC LSS	0.
HUVEC Stat	0.
HUVEC LSS	0.
HUVEC T6	0.
HUVEC RDC	0.
HUVEC RDH	0.
HUVEC HuC	0.
HUVEC HuH	0.
Hemangioma	12
Hemangioma	7.

TABLE 5

Expression of 15603 with β2 in the Angiogenesis Panel	
Tissue Type	Expression
ONC 101 Hemangioma	14.9885
ONC 102 Hemangioma	4.6615

TABLE 5-continued

Expression of 15603 with β2 in the Angiogenesis Panel	
Tissue Type	Expression
ONC 103 Hemangioma	5.1902
CHT 1273 Glioblastoma	0.1909
CHT 216 Glioblastoma	0.0318
CHT 501 Glioblastoma	0.1578
NDR 203 Normal Kidney	0.2082
PIT 213 Renal Cell Carcinoma	0
CHT 732 Wilms Tumor	1.4853
CHT 765 Wilms Tumor	4.4871
NDR 295 Skin	0
CHT 1424 Uterine Adenocarcinoma	1.0911
CHT 1238 Neuroblastoma	0
BWH 78 Fetal Adrenal	0.0346
BWH 74 Fetal Kidney	0.2635
BWH 4 Fetal Heart	0.1023
MPI 849 Normal Heart	0
CLN 796 Spinal Cord	0

Human 69318

[0104] There are over 30 families of secondary transporters, also known as solute carriers or SLC (reviewed by Berger, et al. (2000) in *The Kidney: Physiology and Pathophysiology*, eds. Seldin D W and Giebisch G., Lippincott, Williams & Wilkins, Philadelphia 1:107-138). The SLC families are classified according to the pair of molecules they move. The SLC8 and SLC24 families transport calcium. The SLC8 family members are calcium/sodium antiporters, while the SLC24 family members couple potassium with calcium in exchange for sodium. 69318 is a member of the SLC8 and SLC24 families.

[0105] The human 69318 sequence (SEQ ID NO:13), which is approximately 2875 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 1752 nucleotides, not including the termination codon (nucleotides 115-1866 of SEQ ID NO:13; 1-1752 SEQ ID NO:15). The coding sequence encodes a 584 amino acid protein (SEQ ID NO:14).

[0106] Human 69318 contains the following regions or other structural features (for general information regarding PFAM identifiers, PS prefix and PF prefix domain identification numbers, refer to Sonnhammer et al. (1997) *Protein* 28:405-420: two sodium/calcium exchanger domains (SEQ ID NO:16, PFAM PF01699) located at about amino acid residues 113 to 252 and 431 to 576 of SEQ ID NO:14; twelve transmembrane domains (predicted by MEMSAT, Jones et al. (1994) *Biochemistry* 33:3038-3049) at about amino acids 10 to 26, 95 to 119, 139 to 161, 168 to 192, 205 to 221, 228 to 251, 384 to 408, 419 to 437, 447 to 465, 486 to 509, 525 to 547, and 555 to 573 of SEQ ID NO:14; two protein kinase C phosphorylation sites (Prosite PS00005) at about amino acids 62 to 64 and 132 to 134 of SEQ ID NO:14; four casein kinase II phosphorylation sites (Prosite PS00006) located at about amino acids 70 to 73, 271 to 274, 468 to 471, and 514 to 517 of SEQ ID NO:14; two cAMP/cGMP-dependent protein kinase phosphorylation sites (Prosite PS00004) located at about amino acids 255 to 258 and 318 to 321 of SEQ ID NO:14; two N-glycosylation sites (Prosite PS00001) from about amino acids 60 to 63, and 125 to 128 of SEQ ID NO:14; one amidation site (Prosite PS00009) from about amino acids 2 to 5 of SEQ ID NO:14; and nine N-myristoylation sites (Prosite PS00008) from about amino acids 26 to 31, 58 to 63,

115 to 120, 168 to 173, 178 to 183, 398 to 403, 466 to 471, 494 to 499, and 502 to 507 of SEQ ID NO:14.

[0107] A hydropathy plot of human 69318 was performed. Polypeptides of the invention include fragments which include: all or part of a hydrophobic sequence, e.g., the sequence from about amino acid 10 to 26, from about 168 to 192, and from about 486 to 509 of SEQ ID NO:14; all or part of a hydrophilic sequence, e.g., the sequence from about amino acid 251 to 258, from about 271 to 281, and from about 346 to 357 of SEQ ID NO:14; a sequence which includes a Cys, or a glycosylation site.

[0108] The 69318 protein contains a significant number of structural characteristics in common with transporters, more specifically, with members of the sodium/calcium exchanger families, SLC8 and SLC24. The term "family" when referring to the protein and nucleic acid molecules of the invention means two or more proteins or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein. As used herein, the term "transporter" or "sodium/calcium exchanger" refers to secondary active transport proteins. Secondary active transporters typically couple the active transport of one molecule, e.g., an ion, e.g., a calcium ion against its concentration gradient to the energy gained by concomitant transport of a second molecule, e.g., another ion (e.g., a sodium ion) with its concentration gradient. Human sodium/calcium exchangers have been grouped into two families, named SLC8 and SLC24. In the SLC8 family, the calcium transport is coupled with sodium exchange. In the SLC24 family, the sodium exchange energy is supplemented with additional energy derived from exchange of potassium down its gradient. Thus, in the SLC24 family, calcium moves against its concentration gradient in the same direction as potassium, which moves with its concentration gradient, both at the same time as the opposite movement of sodium with its concentration gradient.

[0109] Typically, sodium/calcium exchangers or SLC8 or SLC24 family members are integral membrane proteins having at least one, two, three, four, five, six, seven, eight, nine, ten, eleven and preferably twelve transmembrane domains. These transmembrane domains can be divided into two homologous groups, one encompassing transmembrane domains 2 to 6 and the other encompassing transmembrane domains 8 to 12. Each group is named a sodium/calcium exchanger domain and is involved in the actual cross-membrane ion transfer. The loops before and between the sodium/calcium exchanger domains are hypervariable and involved in the tissue and ion specificity, ion binding and transporter regulation. The first hypervariable loop is extracellular and the second hypervariable loop is cytoplasmic (Prinsen et al. (2000) *J. Neuroscience* 20:1424-34). In SLC8, the cytoplasmic hypervariable loop is at least 500 amino acids long (Berger et al. supra). SLC24 family members typically have shorter cytoplasmic loops. A GAP alignment of 69318 with human NCKX2, an SLC24 family member (Accession number 6650379 in GenPept, corresponding to AF097366 in Genbank, SEQ ID NO:17 found a 24% identity (as determined using a matrix made by matblas from blosum62.iij).

[0110] A 69318 polypeptide can include at least one, preferably two "sodium/calcium exchanger domains" or regions homologous with a "sodium/calcium exchanger domain".

[0111] As used herein, the term "sodium/calcium exchanger domain" includes an amino acid sequence of about 50 to 250 amino acid residues in length and having a bit score for the alignment of the sequence to the sodium/calcium exchanger domain (HMM) of at least 50. Preferably, a sodium/calcium exchanger domain mediates transport of an ion e.g. a sodium, calcium or potassium ion from one side of a membrane to the opposite side of the membrane. Preferably, a sodium/calcium exchanger domain includes at least about 80 to 200 amino acids, more preferably about 110 to 175 amino acid residues, or about 135 to 150 amino acids and has a bit score for the alignment of the sequence to the sodium/calcium exchanger domain (HMM) of at least 60, 70, 80 or greater. The sodium/calcium exchanger domain (HMM) has been assigned the PFAM Accession Number PF01699. An alignment of the first sodium/calcium exchanger domain (amino acids 113 to 252 of SEQ ID NO:14) of human 69318 with a consensus amino acid sequence (SEQ ID NO:16) derived from a hidden Markov model yields a bit score of 85.2. An alignment of the second sodium/calcium exchanger domain (amino acids 431 to 576 of SEQ ID NO:14) of human 69318 with a consensus amino acid sequence (SEQ ID NO:16) derived from a hidden Markov model yields a bit score of 92.1.

[0112] In a preferred embodiment, a 69318 polypeptide or protein has a "sodium/calcium exchanger domain" or a region which includes at least about 80 to 200, more preferably about 110 to 175 or 135 to 150 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "sodium/calcium exchanger domain," e.g., the sodium/calcium exchanger domain of human 69318 (e.g., residues 113 to 252 and 431 to 576 of SEQ ID NO:14).

[0113] To identify the presence of a "sodium/calcium exchanger" domain in a 69318 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against the Pfam database of HMMs (e.g., the Pfam database, release 2.1) using the default parameters. For example, the hmmsf program, which is available as part of the HMMER package of search programs, is a family specific default program for MILPAT0063 and a score of 15 is the default threshold score for determining a hit. Alternatively, the threshold score for determining a hit can be lowered (e.g., to 8 bits). A description of the Pfam database can be found in Sonhammer et al. (1997) *Proteins* 28:405-420 and a detailed description of HMMs can be found, for example, in Gribskov et al. (1990) *Meth. Enzymol.* 183:146-159; Gribskov et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:4355-4358; Krogh et al. (1994) *J. Mol. Biol.* 235:1501-1531; and Stultz et al. (1993) *Protein Sci.* 2:305-314, the contents of which are incorporated herein by reference. A search was performed against the HMM database resulting in the identification of a "sodium/calcium exchanger domain" domain in the amino acid sequence of human 69318 at about residues 113 to 252 and 431 to 576 of SEQ ID NO:14.

[0114] A 69318 polypeptide can include at least one, two, three, four, five, six, seven, eight, nine, ten, eleven and preferably twelve transmembrane domains or regions homologous with a "transmembrane domain". As used herein, the term "transmembrane domain" includes an amino acid sequence of about 10 to 40 amino acid residues in length and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, e.g., at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, e.g., leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains typically have alpha-helical structures and are described in, for

example, Zagotta, W. N. et al., (1996) *Annual Rev. Neurosci.* 19:235-263, the contents of which are incorporated herein by reference.

[0115] In a preferred embodiment, a 69318 polypeptide or protein has at least one, two, three, four, five, six, seven, eight, nine, ten, eleven and preferably twelve "transmembrane domains" or regions which include at least about 12 to 35 more preferably about 14 to 30 or 15 to 25 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "transmembrane domain," e.g., the transmembrane domains of human 69318 (e.g., residues 10 to 26, 95 to 119, 139 to 161, 168 to 192, 205 to 221, 228 to 251, 384 to 408, 419 to 437, 447 to 465, 486 to 509, 525 to 547, and 555 to 573 of SEQ ID NO:14). The transmembrane domains of human 69318 can be visualized in a hydropathy plot as regions of about 15 to 25 amino acids where the hydropathy trace is mostly above the horizontal line.

[0116] To identify the presence of a "transmembrane" domain in a 69318 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be analyzed by a transmembrane prediction method that predicts the secondary structure and topology of integral membrane proteins based on the recognition of topological models (MEMSAT, Jones et al., (1994) *Biochemistry* 33:3038-3049).

[0117] A 69318 polypeptide can include at least one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve and preferably thirteen "non-transmembrane regions." As used herein, the term "non-transmembrane region" includes an amino acid sequence not identified as a transmembrane domain. The non-transmembrane regions in 69318 are located at about amino acids 1 to 9, 27 to 94, 120 to 138, 162 to 167, 193 to 204, 222 to 227, 252 to 383, 409 to 418, 438 to 446, 466 to 485, 510 to 524, 548 to 554 and 574 to 584 of SEQ ID NO:14.

[0118] In one embodiment, a 69318 protein includes at least one, two, three, four, preferably five cytoplasmic loops. As used herein, the term "loop" includes an amino acid sequence that resides outside of a phospholipid membrane, having a length of at least about 4, preferably about 5 to 600, and more preferably about 5 to 150 amino acid residues, and has an amino acid sequence that connects two transmembrane domains within a protein or polypeptide. As used herein, a "cytoplasmic loop" includes an amino acid sequence having about 1 to 600, preferably about 1 to 400, preferably about 1 to 300, more preferably about 1 to 200, more preferably about 1 to 150, or even more preferably about 1 to 135 amino acid residues in length and is located inside of a cell or intracellularly. The C-terminal amino acid residue of a "cytoplasmic loop" is adjacent to an N-terminal amino acid residue of a transmembrane domain in a 69318 protein. For example, a cytoplasmic loop is located at about amino acid residues 120 to 138, 193 to 204, 252 to 383, 438 to 446, or 510 to 524 of SEQ ID NO:14.

[0119] In a preferred embodiment, a 69318 polypeptide or protein has at least one cytoplasmic loop or a region which includes at least about 5, preferably about 7 to 300, and more preferably about 9 to 150 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "cytoplasmic loop," e.g., at least one cytoplasmic loop of human 69318 (e.g., residues 252 to 383 of SEQ ID NO:14).

[0120] In another embodiment, a 69318 protein includes at least one, two, three, four, five, preferably six non-cytoplasmic loops. As used herein, a "non-cytoplasmic loop" includes an amino acid sequence located outside of a cell or within an intracellular organelle. Non-cytoplasmic loops include extracellular domains (i.e., outside of the cell) and intracellular domains (i.e., within the cell). When referring to membrane-bound proteins found in intracellular organelles (e.g., mitochondria, endoplasmic reticulum, peroxisomes microsomes, vesicles, endosomes, and lysosomes), non-cytoplasmic loops include those domains of the protein that reside in the lumen of the organelle or the matrix or the intermembrane space. Accordingly, the N-terminal amino acid of a non-cytoplasmic loop is adjacent to a C-terminal amino acid of a transmembrane domain in a 69318 molecule, and the C-terminal amino acid of a non-cytoplasmic loop is adjacent to an N-terminal amino acid of a transmembrane domain in a 69318 molecule. For example, a "non-cytoplasmic loop" can be found at about amino acids 27 to 94, 162 to 167, 222 to 227, 409 to 418, 466 to 485, and 548 to 554 of SEQ ID NO:14.

[0121] In a preferred embodiment, a 69318 polypeptide or protein has at least one non-cytoplasmic loop or a region which includes at least about 4, preferably about 5 to 40, preferably about 6 to 60, and more preferably about 6 to 70 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with an "non-cytoplasmic loop," e.g., at least one non-cytoplasmic loop of human 69318 (e.g., residues 27 to 94 of SEQ ID NO:14).

[0122] A 69318 family member can include at least one, preferably two sodium/calcium exchanger domains; and at least one, two, three, four, five, six, seven, eight, nine, ten, eleven and preferably twelve transmembrane domains; at least one cytoplasmic loop; at least one non-cytoplasmic loop; or at least one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve and preferably thirteen non-transmembrane regions. Furthermore, a 69318 family member can include at least one, preferably two protein kinase C phosphorylation sites (PS00005); at least one, two, three, and preferably four casein kinase II phosphorylation sites (PS00006); at least one, preferably two N-glycosylation sites (PS00001); at least one, preferably two cAMP- and cGMP-dependent protein kinase phosphorylation sites (PS00004); at least one amidation site (PS00009); and at least one, three, five, seven and preferably nine N-myristoylation sites (PS00008).

[0123] As the 69318 polypeptides of the invention can modulate 69318-mediated activities, they can be useful for developing novel diagnostic and therapeutic agents for sodium/calcium exchanger-associated or other 69318-associated disorders, as described below.

[0124] Sodium/calcium exchangers actively transport calcium and participate in calcium homeostasis. These transporters can be found in the heart, muscle, brain, retina and kidney, especially in the excitable cells, where they participate in the many physiological processes which require management of intracellular calcium levels. These transporters dynamically coordinate with calcium channels and calcium binding proteins to control the availability of these ions for calcium-dependent cellular responses. For example, in an excitable or contractile cell, a stimulus activates calcium channels to allow rapid cytosolic influx of calcium and induce a response to the stimulus. After the stimulus is removed, the sodium/calcium exchangers transport the calcium back out of the cytoplasm to restore the potential function of the cell.

[0125] Alterations in the dynamic interplay of calcium influx and efflux processes play roles in many diseases. These diseases can be associated with an intracellular calcium over-

load, such as in MELAS, an encephalomyopathy (Moudy et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:729-33), in retinal degeneration (Edward et al. (1991) *Arch. Opthalmol.* 109: 554-62), or in myocardial reoxygenation/reperfusion injury (Mochizuki and Jiang (1998) *Jpn. Heart J.* 39:707-14); an intracellular potassium overload in a variant of Bartter's syndrome (Peleg et al. (1997) *Hypertension* 30:1338-41); or an intracellular calcium deficiency in stimulated platelets of some diabetic patients (Yamaguchi et al. (1991) *Diabetes Res.* 18:89-94).

[0126] As used herein, a "69318 activity", "biological activity of 69318" or "functional activity of 69318", refers to an activity exerted by a 69318 protein, polypeptide or nucleic acid molecule on e.g., a 69318-responsive cell or on a 69318 substrate, e.g., a protein substrate, as determined in vivo or in vitro. In one embodiment, a 69318 activity is a direct activity, such as an association with a 69318 target molecule. A "target molecule" or "binding partner" is a molecule with which a 69318 protein binds or interacts in nature. In an exemplary embodiment, 69318 is a transporter, e.g., an SLC8 family sodium/calcium exchanger or SLC24 family sodium/calcium/potassium exchanger, and thus binds to or interacts in nature with a molecule, e.g., an ion, (e.g., a calcium ion), a second molecule, e.g., an ion, (e.g., a sodium ion), and/or a third molecule, e.g., an ion, (e.g., a potassium ion).

[0127] A 69318 activity can also be an indirect activity, e.g., a cellular signaling activity mediated by interaction of the 69318 protein with a 69318 receptor. Based on the above-described sequence structures and similarities to molecules of known function, the 69318 molecules of the present invention have similar biological activities as sodium/calcium exchanger family members. For example, the 69318 proteins of the present invention can have one or more of the following activities: (1) the ability to reside within a membrane, e.g., a cell or organelle membrane; (2) the ability to interact with, e.g., bind to, a substrate or target molecule; (3) the ability to transport a substrate or target molecule, e.g., an ion (e.g., a calcium ion) across a membrane; (4) the ability to transport a second substrate or target molecule, e.g., another ion (e.g., a sodium ion) across a membrane; (5) the ability to transport a third substrate or target molecule, e.g., another ion (e.g., a potassium ion) across a membrane; (6) the ability to interact with and/or modulate the activity of a second non-transporter protein; (7) the ability to modulate cellular signaling and/or gene transcription (e.g., either directly or indirectly); (8) the ability to modulate calcium homeostasis; (9) the ability to modulate muscle contraction; (10) the ability to modulate responses to stimuli or (11) the ability to modulate vision.

[0128] The 69318 molecules of the invention can modulate the activities of cells in tissues where they are expressed. TaqMan analysis shows 69318 mRNA is expressed in normal artery, human umbilical vein endothelial cells (HUVEC), kidney, pancreas, normal brain cortex, breast tumor, normal ovary, and lung tumor. Accordingly, the 69318 molecules of the invention can act as therapeutic or diagnostic agents for cardiovascular disorders, including endothelial cell disorders, renal disorders, pancreatic disorders, neurological disorders, cellular proliferative and/or differentiative disorders and ovarian disorders.

[0129] The 69318 molecules can be used to treat cardiovascular disorders in part because the 69318 mRNA is expressed in normal artery, and human umbilical vein endothelial cells. A cardiovascular disease or disorder also includes an endothelial cell disorder.

[0130] The 69318 molecules can be used to treat renal disorders in part because the 69318 mRNA is expressed in the kidney.

[0131] The 69318 molecules can be used to treat pancreatic disorders in part because the 69318 mRNA is expressed in the pancreas.

[0132] The 69318 molecules can be used to treat neurological disorders in part because the 69318 mRNA is expressed in the brain cortex

[0133] The 69318 molecules can be used to treat cellular proliferative and/or differentiative disorders in part because the 69318 mRNA is expressed in lung tumor and breast tumor but not in normal lung or normal breast.

[0134] The 69318 molecules can be used to treat ovarian disorders in part because the 69318 mRNA is expressed in normal ovary.

[0135] Thus, the 69318 molecules can act as novel diagnostic targets and therapeutic agents for controlling one or more cardiovascular disorders, including endothelial cell disorders, renal disorders, pancreatic disorders, neurological disorders, and cellular proliferative and differentiative disorders and other transporter, e.g., sodium/calcium exchanger or sodium/calcium/potassium exchanger disorders. Examples of such disorders, e.g., sodium/calcium exchanger-associated or other 69318-associated disorders, include but are not limited to, eye and vision disorders, immune and inflammatory disorders, hematopoietic disorders, pain disorders, or metabolic disorders.

[0136] The 69318 nucleic acid and protein of the invention can be used to treat and/or diagnose a variety of immune, e.g., inflammatory disorders.

[0137] The 69318 molecules of the invention can be used to monitor, treat and/or diagnose a variety of proliferative disorders. Such disorders include hematopoietic neoplastic disorders.

[0138] Additionally, 69318 can play an important role in the regulation of metabolism or pain disorders.

Gene Expression Analysis of 69318

[0139] Human 69318 expression was measured by TaqMan® quantitative PCR (Perkin Elmer Applied Biosystems) in cDNA prepared from a variety of normal and diseased (e.g., cancerous) human tissues or cell lines.

[0140] The results indicate significant 69318 expression in normal artery, human umbilical vein endothelial cells (HUVEC), kidney, pancreas, normal brain cortex, breast tumor, normal ovary, and lung tumor.

Human 12303 (TWIK-8)

[0141] The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as "TWIK-8" or "12303" nucleic acid and protein molecules, which are novel members of the TWIK (for Tandem of P domains in a Weak Inward rectifying K^+ channel)-like family of potassium channels. These novel molecules are capable of, for example, modulating a potassium channel mediated activity in a cell, e.g., a neuronal cell, or a muscle cell.

[0142] Potassium (K^+) channels are ubiquitous proteins which are involved in the setting of the resting membrane potential as well as in the modulation of the electrical activity of cells. In excitable cells, K^+ channels influence action potential waveforms, firing frequency, and neurotransmitter secretion (Rudy, B. (1988) *Neuroscience,* 25, 729-749; Hille,

B. (1992) *Ionic Channels of Excitable Membranes,* 2nd Ed.). In non-excitable cells, they are involved in hormone secretion, cell volume regulation and potentially in cell proliferation and differentiation (Lewis et al. (1995) *Annu. Rev. Immunol.,* 13, 623-653). Developments in electrophysiology have allowed the identification and the characterization of an astonishing variety of K^+ channels that differ in their biophysical properties, pharmacology, regulation and tissue distribution (Rudy, B. (1988) *Neuroscience,* 25, 729-749; Hille, B. (1992) *Ionic Channels of Excitable Membranes,* 2nd Ed.). More recently, cloning efforts have shed considerable light on the mechanisms that determine this functional diversity. Furthermore, analyses of structure-function relationships have provided an important set of data concerning the molecular basis of the biophysical properties (selectivity, gating, assembly) and the pharmacological properties of cloned K^+ channels.

[0143] Functional diversity of K^+ channels arises mainly from the existence of a great number of genes coding for pore-forming subunits, as well as for other associated regulatory subunits. Two main structural families of pore-forming subunits have been identified. The first one consists of subunits with a conserved hydrophobic core containing six transmembrane domains (TMDs). These K^+ channel α subunits participate in the formation of outward rectifier voltage-gated (Kv) and Ca^{2+}-dependent K^+ channels. The fourth TMD contains repeated positive charges involved in the voltage gating of these channels and hence in their outward rectification (Logothetis et al. (1992) *Neuron,* 8, 531-540; Bezanilla et al. (1994) *Biophys. J.* 66, 1011-1021).

[0144] The second family of pore-forming subunits have only two TMDs. They are essential subunits of inward-rectifying (IRK), G-protein-coupled (GIRK) and ATP-sensitive (K_{ATP}) K^+ channels. The inward rectification results from a voltage-dependent block by cytoplasmic Mg^{2+} and polyamines (Matsuda, H. (1991) *Annu. Rev. Physiol.,* 53, 289-298). A conserved domain, called the P domain, is present in all members of both families (Pongs, O. (1993) *J. Membr. Biol.,* 136, 1-8; Heginbotham et al. (1994) *Biophys. J.* 66, 1061-1067; Mackinnon, R. (1995) *Neuron,* 14, 889-892; Pascual et al., (1995) *Neuron*., and 14, 1055-1063). This domain is an essential element of the aqueous K^+-selective pore. In both groups, the assembly of four subunits is necessary to form a functional K^+ channel (Mackinnon, R. (1991) *Nature,* 350, 232-235; Yang et al., (1995) *Neuron,* 15, 1441-1447.

[0145] In both six TMD and two TMD pore-forming subunit families, different subunits coded by different genes can associate to form heterotetramers with new channel properties (Isacoff et al., (1990) *Nature,* 345, 530-534). A selective formation of heteropolymeric channels may allow each cell to develop the best K^+ current repertoire suited to its function. Pore-forming ax subunits of Kv channels are classified into different subfamilies according to their sequence similarity (Chandy et al. (1993) *Trends Pharmacol. Sci.,* 14: 434). Tetramerization is believed to occur preferentially between members of each subgroup (Covarrubias et al. (1991) *Neuron,* 7, 763-773). The domain responsible for this selective association is localized in the N-terminal region and is conserved between members of the same subgroup. This domain is necessary for hetero- but not homo-multimeric assembly within a subfamily and prevents co-assembly between subfamilies. Recently, pore-forming subunits with two TMDs were also shown to co-assemble to form heteropolymers (Duprat et al. (1995) *Biocheni. Biophys. Res. Commun.,* 212, 657-663. This heteropolymerization seems necessary to give functional GIRKs. IRKs are active as homopolymers but also form heteropolymers.

[0146] New structural types of K^+ channels were identified recently in both humans and yeast. These channels have two P domains in their functional subunit instead of only one (Ketchum et al. (1995) *Nature,* 376, 690-695; Lesage et al. (1996) *J. Biol. Chem.,* 271, 4183-4187; Lesage et al. (1996) *EMBO J.,* 15, 1004-1011; Reid et al. (1996) *Receptors Channels* 4, 51-62). The human channel called TWIK-1, has four TMDs. TWIK-1 is expressed widely in human tissues and is particularly abundant in the heart and the brain. TWIK-1 currents are time independent and inwardly rectifying. These properties suggest that TWIK-1 channels are involved in the control of the background K^+ membrane conductance (Lesage et al. (1996) *EMBO J.,* 15, 1004-1011).

[0147] The human TWIK-8 or 12303 sequence (SEQ ID NO:18), which is approximately 1408 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 1257 nucleotides, not including the termination codon (nucleotides 84-1340 of SEQ ID NO:18; 1-1257 SEQ ID NO:20). The coding sequence encodes a 419 amino acid protein (SEQ ID NO:19).

[0148] The amino acid sequence of human TWIK-8 was analyzed using the program PSORT to predict the localization of the proteins within the cell. This program assesses the presence of different targeting and localization amino acid sequences within the query sequence. The results of the analysis show that human TWIK-8 (SEQ ID NO:19) may be localized to the endoplasmic reticulum or to the mitochondrion.

[0149] An analysis of the amino acid sequence of human TWIK-8 using the Signal P program (Henrik, et al. (1997) *Protein Engineering* 10:1-6), identified the presence of a signal peptide from amino acids 1-46 of SEQ ID NO:19.

[0150] A search was performed against the Memsat database resulting in the identification of six transmembrane domains in the amino acid sequence of the native human TWIK-8 (SEQ ID NO: 19) at about residues 32-50, 116-137, 144-165, 195-219, 226-242, and 260-283. This search further identified five transmembrane domains in the amino acid sequence of the predicted mature form of this protein, at about residues 70-91, 98-119, 149-173, 180-196, and 214-237 of SEQ ID NO:19.

[0151] A search was performed against the HMM database, resulting in the identification of a "seven-transmembrane receptor domain" from about residues 25-244, and a "cyclic nucleotide-gated channel domain" from about residues 27-204 in the amino acid sequence of human TWIK-8 (SEQ ID NO:19).

[0152] A search was also performed against the ProDom database, resulting in the identification of "TRAAK potassium channel domains" from about residues 50-104 (score=175), 175-199 (score=115), and 288-382 (score=135) of SEQ ID NO:19; a "potassium channel protein domain" from about residues 99-153 (score=101) of SEQ ID NO:19; a "voltage-gated potassium channel domain" from about residues 102-168 (score=115) of SEQ ID NO:19; an "outward-rectifier TOK1 potassium channel domain" from about residues 215-270 (score=70) of SEQ ID NO:19; and a "potassium channel subunit domain" from about residues 216-287 (score=156) in the amino acid sequence of human TWIK-8 (SEQ ID NO:19).

[0153] A BLASTX 2.0 search against the NRP/protot database, using a score of 100, a wordlength of 3, and a Blosum 62 matrix (Altschul et al. (1990) *J. Mol. Biol.* 215:403), of the translated nucleotide sequence of human TWIK-8 revealed that human TWIK-8 has limited sequence homology to *Mus musculus* TRAAK K+ channel subunit mRNA (Accession Number AF056492), to *Homo sapiens* TREK-1 potassium channel (KCNK2) mRNA (Accession Number AF129399), to *Mus musculus* TREK-1 K+ channel subunit mRNA (Accession Number U73488), and to *Homo sapiens* two-pore potassium channel TPKC1 mRNA (Accession number AF004711).

[0154] As used herein, a "potassium channel" includes a protein or polypeptide which is involved in receiving, conducting, and transmitting signals in an electrically excitable or a non-electrically excitable cell, e.g., a neuronal cell, or a muscle cell (e.g., a cardiac muscle cell). Potassium channels are potassium ion selective, and can determine membrane excitability (the ability of, for example, a neuron to respond to a stimulus and convert it into an impulse). Potassium channels can also influence the resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation. Potassium channels are typically expressed in electrically excitable cells, e.g., neurons, muscle, endocrine, and egg cells, and may form heteromultimeric structures, e.g., composed of pore-forming α and cytoplasmic β subunits. Potassium channels may also be found in non-excitable cells (e.g., spleen cells or prostate cells), where they may play a role in, e.g., signal transduction. Examples of potassium channels include: (1) the voltage-gated potassium channels, (2) the ligand-gated potassium channels, e.g., neurotransmitter-gated potassium channels, and (3) cyclic-nucleotide-gated potassium channels. Voltage-gated and ligand-gated potassium channels are expressed in the brain, e.g., in brainstem monoaminergic and forebrain cholinergic neurons, where they are involved in the release of neurotransmitters, or in the dendrites of hippocampal and neocortical pyramidal cells, where they are involved in the processes of learning and memory formation. For a detailed description of potassium channels, see Kandel E. R. et al., Principles of Neural Science, second edition, (Elsevier Science Publishing Co., Inc., N.Y. (1985)), the contents of which are incorporated herein by reference. As the TWIK-like proteins of the present invention may modulate potassium channel mediated activities, they may be useful for developing novel diagnostic and therapeutic agents for potassium channel associated disorders.

[0155] As used herein, a "potassium channel associated disorder" includes a disorder, disease or condition which is characterized by a misregulation of a potassium channel mediated activity. Potassium channel associated disorders can detrimentally affect conveyance of sensory impulses from the periphery to the brain and/or conductance of motor impulses from the brain to the periphery; integration of reflexes; interpretation of sensory impulses; and emotional, intellectual (e.g., learning and memory), or motor processes.

[0156] Examples of potassium channel associated disorders include CNS disorders such as cognitive and neurodegenerative disorders.

[0157] Further examples of potassium channel associated disorders include cardiac-related disorders. TWIK-8-mediated or related disorders also include disorders of the musculoskeletal system such as paralysis and muscle weakness, e.g., ataxia, myotonia, and myokymia.

[0158] Other examples of potassium channel-associated disorders include pain disorders. Pain disorders include those disorders that affect pain signaling mechanisms. The TWIK-8 molecules of the present invention may be present on these sensory neurons and, thus, may be involved in detecting these noxious chemical, mechanical, or thermal stimuli and transducing this information into membrane depolarization events. Thus, the TWIK-8 molecules by participating in pain signaling mechanisms, may modulate pain elicitation and act as targets for developing novel diagnostic targets and therapeutic agents to control pain.

[0159] Potassium channel disorders also include cellular proliferation, growth, differentiation, or migration disorders. Cellular proliferation, growth, differentiation, or migration disorders include those disorders that affect cell proliferation, growth, differentiation, or migration processes.

[0160] The TWIK-8 molecules of the present invention are involved in signal transduction mechanisms, which are known to be involved in cellular growth, differentiation, and migration processes. Thus, the TWIK-8 molecules may modulate cellular growth, differentiation, or migration, and may play a role in disorders characterized by aberrantly regulated growth, differentiation, or migration. Such disorders include cancer, e.g., carcinoma, sarcoma, or leukemia; tumor angiogenesis and metastasis; skeletal dysplasia; neuronal deficiencies resulting from impaired neural induction and patterning; and hematopoietic and/or myeloproliferative disorders.

[0161] TWIK-8-associated or related disorders also include disorders of tissues in which TWIK-8 protein is expressed, e.g., brain cortex, hypothalamus and dorsal root ganglia.

[0162] As used herein, a "potassium channel mediated activity" includes an activity which involves a potassium channel, e.g., a potassium channel in a neuronal cell, or a muscle cell (e.g., a cardiac muscle cell), associated with receiving, conducting, and transmitting signals in, for example, the nervous system. Potassium channel mediated activities include release of neurotransmitters, e.g., dopamine or norepinephrine, from cells, e.g., neuronal cells; modulation of resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation; and modulation of processes such as integration of subthreshold synaptic responses, the conductance of back-propagating action potentials in, for example, neuronal cells or muscle cells, participation in signal transduction pathways, and participation in nociception.

[0163] The term "family" when referring to the protein and nucleic acid molecules of the invention is intended to mean two or more proteins or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein. Such family members can be naturally or non-naturally occurring and can be from either the same or different species. For example, a family can contain a first protein of human origin, as well as other, distinct proteins of human origin or alternatively, can contain homologues of non-human origin, e.g., monkey proteins. Members of a family may also have common functional characteristics.

[0164] For example, the family of TWIK-8 proteins comprises at least one "transmembrane domain" and preferably six transmembrane domains. As used herein, the term "transmembrane domain" includes an amino acid sequence of about 15 amino acid residues in length which spans the

plasma membrane. More preferably, a transmembrane domain includes about at least 20, 25, 30, 35, 40, or 45 amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an alpha-helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, e.g., leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains are described in, for example, Zagotta W. N. et al. (1996) *Annual Rev. Neurosci.* 19: 235-263, the contents of which are incorporated herein by reference. Amino acid residues 32-50, 116-137, 144-165, 195-219, 226-242, and 260-283 of the native TWIK-8 protein (SEQ ID NO:19), and amino acid residues 70-91, 98-119, 149-173, 180-196 and 214-237 of the mature TWIK-8 protein are predicted to comprise transmembrane domains. Accordingly, TWIK-8 proteins having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with a transmembrane domain of human TWIK-8 are within the scope of the invention.

[0165] In another embodiment, a TWIK-8 molecule of the present invention is identified based on the presence of a Pore loop (P-loop). As used herein, the term "Pore loop" or "P-loop" includes an amino acid sequence of about 15-45 amino acid residues in length, preferably about 15-35 amino acid residues in length, and most preferably about 15-25 amino acid residues in length, which is involved in lining the potassium channel pore. A P-loop is typically found between transmembrane domains of potassium channels and is believed to be a major determinant of ion selectivity in potassium channels. Preferably, P-loops contain a G-[HYDROPHOBIC AMINO ACID]-G sequence, e.g., a GYG, GLG, or GFG sequence. P-loops are described in, for example, Warmke et al. (1991) *Science* 252:1560-1562; Zagotta W. N. et al., (1996) *Annual Rev. Neuronsci.* 19:235-63 (Pongs, O. (1993) *J. Membr. Biol.,* 136, 1-8; Heginbotham et al. (1994) *Biophys. J.* 66, 1061-1067; Mackinnon, R. (1995) *Neuron, and* 14, 889-892; Pascual et al., (1995) *Neuron.,* 14, 1055-1063), the contents of which are incorporated herein by reference. Amino acid residues 243-259 (SEQ ID NO:19) of the native human TWIK-8 protein, and residues 197-213 of the predicted mature human TWIK-8 protein comprise a P-loop.

[0166] In a preferred embodiment, the TWIK-8 molecules of the invention include at least one and, preferably, six transmembrane domains and at least one P-loop domain.

[0167] In another embodiment, a TWIK-8 molecule of the present invention is identified based on the presence of a "seven-transmembrane receptor domain" in the protein or corresponding nucleic acid molecule. Seven-transmembrane receptor domains are described, for example, in Hamann et al. (1996) *Genomics* 32: 144-147. As used herein, the term "seven-transmembrane receptor domain" includes a protein domain having an amino acid sequence of about 150-320 amino acid residues. Preferably, a seven-transmembrane receptor domain includes at least about 200-250, or more preferably about 220 amino acid residues. To identify the presence of a seven-transmembrane receptor domain in a TWIK-8 protein, and to make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein is searched against a database of known protein domains (e.g., the HMM database). The seven-transmembrane receptor domain (HMM) has been assigned the PFAM Accession PF00002. A search was performed against the HMM database resulting in the identification of a seven-transmembrane receptor domain in the amino acid sequence of human TWIK-8 at about residues 25-244 of SEQ ID NO: 19.

[0168] In another embodiment, a TWIK-8 molecule of the present invention is identified based on the presence of a "cyclic nucleotide-gated channel domain" in the protein or corresponding nucleic acid molecule. Cyclic nucleotide-gated channel domains are described, for example, in Yau (1994) *Proc. Natl. Acad. Sci. USA* 91: 3481-3483. As used herein, the term "cyclic nucleotide-gated channel domain" includes a protein domain having an amino acid sequence of about 100-225 amino acid residues. Preferably, a cyclic nucleotide-gated channel domain includes at least about 150-200, or more preferably about 178 amino acid residues. To identify the presence of a cyclic nucleotide-gated channel domain in a TWIK-8 protein, and to make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein is searched against a database of known protein domains (e.g., the HMM database). The cyclic nucleotide-gated channel domain (HMM) has been assigned the PFAM Accession PF00914. A search was performed against the HMM database resulting in the identification of a cyclic nucleotide-gated channel domain in the amino acid sequence of human TWIK-8 at about residues 27-204 of SEQ ID NO:19.

[0169] In another embodiment, a TWIK-8 molecule of the present invention is identified based on the presence of a "TRAAK potassium channel domain" in the protein or corresponding nucleic acid molecule. As used herein, the term "TRAAK potassium channel domain" includes a protein domain having an amino acid sequence of about 20-150 amino acid residues and having a bit score for the alignment of the sequence to the TRAAK potassium channel domain of at least 115-175. Preferably, a TRAAK potassium channel domain includes at least about 23-100, or more preferably about 25, 55, or 95 amino acid residues, and has a bit score for the alignment of the sequence to the TRAAK potassium channel domain of at least 20, 30, 40, 50, 60, 70, 80, 90, 100, 110 or higher. The TRAAK potassium channel domain has been assigned ProDom entries 73512, 98483, and 105542. To identify the presence of a TRAAK potassium channel domain in a TWIK-8 protein, and to make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein is searched against a database of known protein domains (e.g., the ProDom database) using the default parameters. A search was performed against the ProDom database resulting in the identification of a TRAAK potassium channel domains in the amino acid sequence of human TWIK-8 at about residues 50-104, 175-199, and 288-382 of SEQ ID NO:19.

[0170] In another embodiment, a TWIK-8 molecule of the present invention is identified based on the presence of a "potassium channel protein domain" in the protein or corresponding nucleic acid molecule. As used herein, the term "potassium channel protein domain" includes a protein domain having an amino acid sequence of about 20-100 amino acid residues and having a bit score for the alignment of the sequence to the potassium channel protein domain of at least 101. Preferably, a potassium channel protein domain includes at least about 40-75, or more preferably about 55 amino acid residues, and has a bit score for the alignment of the sequence to the potassium channel protein domain of at least 20, 30, 40, 50, 60, 70, 80, 90, 100, or higher. The potassium channel protein domain has been assigned Pro-

Dom entry 129403. To identify the presence of a potassium channel protein domain in a TWIK-8 protein, and to make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein is searched against a database of known protein domains (e.g., the ProDom database) using the default parameters. A search was performed against the ProDom database resulting in the identification of a potassium channel protein domain in the amino acid sequence of human TWIK-8 at about residues 99-153 of SEQ ID NO:19.

[0171] In another embodiment, a TWIK-8 molecule of the present invention is identified based on the presence of a "voltage-gated potassium channel domain" in the protein or corresponding nucleic acid molecule. As used herein, the term "voltage-gated potassium channel domain" includes a protein domain having an amino acid sequence of about 20-100 amino acid residues and having a bit score for the alignment of the sequence to the potassium channel protein domain of at least 115. Preferably, a voltage-gated potassium channel domain includes at least about 40-75, or more preferably about 55 amino acid residues, and has a bit score for the alignment of the sequence to the voltage-gated potassium channel domain of at least 20, 30, 40, 50, 60, 70, 80, 90, 100, 110 or higher. The voltage-gated potassium channel domain has been assigned ProDom entry 36. To identify the presence of a voltage-gated potassium channel domain in a TWIK-8 protein, and to make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein is searched against a database of known protein domains (e.g., the ProDom database) using the default parameters. A search was performed against the ProDom database resulting in the identification of a voltage-gated potassium channel domain in the amino acid sequence of human TWIK-8 at about residues 102-168 of SEQ ID NO:19.

[0172] In another embodiment, a TWIK-8 molecule of the present invention is identified based on the presence of an "outward-rectifier TOK1 potassium channel domain" in the protein or corresponding nucleic acid molecule. As used herein, the term "outward-rectifier TOK1 potassium channel domain" includes a protein domain having an amino acid sequence of about 25-100 amino acid residues and having a bit score for the alignment of the sequence to the outward-rectifier TOK1 potassium channel domain of at least 70. Preferably, an outward-rectifier TOK1 potassium channel domain includes at least about 40-75, or more preferably about 56 amino acid residues, and has a bit score for the alignment of the sequence to the outward-rectifier TOK1 potassium channel domain of at least 20, 30, 40, 50, 60, or higher. The outward-rectifier TOK1 potassium channel domain has been assigned ProDom entry 32818. To identify the presence of an outward-rectifier TOK1 potassium channel domain in a TWIK-8 protein, and to make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein is searched against a database of known protein domains (e.g., the ProDom database) using the default parameters. A search was performed against the ProDom database resulting in the identification of an outward-rectifier TOK1 potassium channel domain in the amino acid sequence of human TWIK-8 at about residues 215-270 of SEQ ID NO:19.

[0173] In another embodiment, a TWIK-8 molecule of the present invention is identified based on the presence of a "potassium channel subunit domain" in the protein or corresponding nucleic acid molecule. As used herein, the term "potassium channel subunit domain" includes a protein domain having an amino acid sequence of about 25-125 amino acid residues and having a bit score for the alignment of the sequence to the potassium channel subunit domain of at least 156. Preferably, a potassium channel subunit domain includes at least about 40-100, or more preferably about 72 amino acid residues, and has a bit score for the alignment of the sequence to the potassium channel subunit domain of at least 20, 30, 40, 50, 60, or higher. The potassium channel subunit domain has been assigned ProDom entry 1641. To identify the presence of a potassium channel subunit domain in a TWIK-8 protein, and to make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein is searched against a database of known protein domains (e.g., the ProDom database) using the default parameters. A search was performed against the ProDom database resulting in the identification of a potassium channel subunit domain in the amino acid sequence of human TWIK-8 at about residues 216-287 of SEQ ID NO:19.

[0174] Isolated proteins of the present invention, preferably TWIK-8 proteins, have an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:19 or are encoded by a nucleotide sequence sufficiently identical to SEQ ID NO:18 or 20. As used herein, the term "sufficiently identical" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (e.g., an amino acid residue which has a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences share common structural domains or motifs and/or a common functional activity. For example, amino acid or nucleotide sequences which share common structural domains have at least 30%, 40%, or 50% homology, preferably 60% homology, more preferably 70%-80%, and even more preferably 90-95% homology across the amino acid sequences of the domains and contain at least one and preferably two structural domains or motifs, are defined herein as sufficiently identical. Furthermore, amino acid or nucleotide sequences which share at least 30%, 40%, or 50%, preferably 60%, more preferably, 70-80%, or 90-95% homology and share a common functional activity are defined herein as sufficiently identical.

[0175] As used interchangeably herein, an "TWIK-8 activity", "biological activity of TWIK-8" or "functional activity of TWIK-8", refers to an activity exerted by a TWIK-8 protein, polypeptide or nucleic acid molecule on a TWIK-8 responsive cell or tissue, or on a TWIK-8 protein substrate, as determined in vivo, or in vitro, according to standard techniques. In one embodiment, a TWIK-8 activity is a direct activity, such as an association with a TWIK-8-target molecule. As used herein, a "target molecule" or "binding partner" is a molecule with which a TWIK-8 protein binds or interacts in nature, such that TWIK-8-mediated function is achieved. A TWIK-8 target molecule can be a non-TWIK-8 molecule or a TWIK-8 protein or polypeptide of the present invention. In an exemplary embodiment, a TWIK-8 target molecule is a TWIK-8 ligand, e.g., a potassium channel pore-forming subunit or a potassium channel ligand. Alternatively, a TWIK-8 activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the TWIK-8 protein with a TWIK-8 ligand. The biological activities of TWIK-8 are described herein. For example, the TWIK-8 proteins of the present invention can have one or more of the following activities: (1) interacting with a non-TWIK protein

molecule; (2) activating a TWIK-dependent signal transduction pathway; (3) modulating the release of neurotransmitters; (4) modulating membrane excitability; (5) influencing the resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation, (6) modulating processes which underlie learning and memory, such as integration of sub-threshold synaptic responses and the conductance of back-propagating action potentials, and (7) mediating nociception.

[0176] Accordingly, another embodiment of the invention features isolated TWIK-8 proteins and polypeptides having a TWIK-8 activity. Preferred proteins are TWIK-8 proteins having at least one or more of the following domains: a transmembrane domain, a pore loop domain, a seven-transmembrane receptor domain, a cyclic nucleotide-gated channel domain, a TRAAK potassium channel domain, a potassium channel protein domain, a voltage-gated potassium channel domain, a potassium channel subunit domain, and an outward-rectifier TOK1 potassium channel domain, and, preferably, a TWIK-8 activity.

[0177] Additional preferred proteins have at least one or more of the following domains: a transmembrane domain, a pore loop domain, a seven-transmembrane receptor domain, a cyclic nucleotide-gated channel domain, a TRAAK potassium channel domain, a potassium channel protein domain, a voltage-gated potassium channel domain, a potassium channel subunit domain, and an outward-rectifier TOK1 potassium channel domain, and are, preferably, encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:18 or 20.

Tissue Distribution of Human TWIK-8 mRNA Using Taqman™ Analysis

[0178] The tissue distribution of human TWIK-8 mRNA in a variety of cells and tissues was determined using the TaqMan™ procedure.

[0179] Highest expression of TWIK-8 mRNA was detected in brain cortex, followed by dorsal root ganglia, and hypothalamus. Weak expression was also detected in erythroid tissue, followed by HUVEC, spinal cord, hemangioma, kidney, normal ovary and ovary tumor, megakaryocytes, normal prostate and prostate tumor. Weak expression was also detected in breast tumor although no expression was detected in normal breast tissue.

Human 48000 (TLCC-4) and 52920 (TLCC-5)

[0180] The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as "TRP-like calcium channel-4" and "TRP-like calcium channel-5" or "TLCC-4" or "48000" and "TLCC-5" or "52920" nucleic acid and polypeptide molecules, which are novel members of the ion channel, e.g., calcium channel and/or vanilloid receptor, family. These novel molecules are capable of, for example, modulating an ion-channel mediated activity (e.g., a calcium channel- and/or vanilloid receptor-mediated activity) in a cell, e.g., a neuronal, skin, muscle (e.g., cardiac muscle), or liver cell.

[0181] Calcium signaling has been implicated in the regulation of a variety of cellular responses, such as growth and differentiation. There are two general methods by which intracellular concentrations of calcium ions may be increased: calcium ions may be freed from intracellular stores, transported by specific membrane channels in the storage organelle, or calcium ions may be brought into the cell from the extracellular milieu through the use of specific channels in the cellular membrane. In the situation in which the intracellular stores of calcium have been depleted, a specific type of calcium channel, termed a 'capacitative calcium channel' or a 'store-operated calcium channel' (SOC), is activated in the plasma membrane to import calcium ions from the extracellular environment to the cytosol (see Putney and McKay (1999) *BioEssays* 21:38-46). Calcium may also enter the cell via receptor-stimulated cation channels (see Hofmann et al. (2000) *J. Mol. Med.* 78:14-25).

[0182] There is no single electrophysiological profile characteristic of the calcium channel family; rather, a wide array of single channel conductances, cation selectivity, and current properties have been observed for different channels. Further, in several instances it has been demonstrated that homo- or hetero-polymerization of the channel molecule may occur, further changing the channel properties from those of the single molecule. In general, though, these channels function similarly, in that they are calcium ion-permeable cation channels which become activated after agonist binding to a G protein-coupled receptor.

[0183] Members of the capacitative calcium channel family include the calcium release-activated calcium current (CRAC) (Hoth and Penner (1992) *Nature* 355: 353-355), calcium release-activated non-selective cation current (CRANC) (Krause et al. (1996) *J. Biol. Chem.* 271: 32523-32528), and the transient receptor potential (TRP) proteins TRP1, TRP2, TRP4, and TRP5. Depletion of intracellular calcium stores activate these channels by a mechanism which is yet undefined, but which has been demonstrated to involve a diffusible factor using studies in which calcium stores were artificially depleted (e.g., by the introduction of chelators into the cell, by activating phospholipase C_γ, or by inhibiting those enzymes responsible for pumping calcium ions into the stores or those enzymes responsible for maintaining resting intracellular calcium ion concentrations) (Putney, J. W., (1986) *Cell Calcium* 7:1-12; Putney, J. W. (1990) *Cell Calcium* 11:611-624).

[0184] Recently, it has been elucidated that three TRP family members, TRP3, TRP6, and a mouse homologue, TRP7, form a sub-family of receptors that are activated in a calcium store-depletion independent manner. TRP3 and TRP6 are activated by diacylglycerols in a membrane delimited manner (Hofmann et al. (1999) *Nature* 397:259-263). Similarly, murine TRP7 is activated via diacylglycerol stimulation by G_q protein coupled receptors (Okada et al. (1999) *J. Biol. Chem.* 274:27359-27370).

[0185] The TRP channel family is one of the best characterized calcium channel protein families. These channels include transient receptor potential protein and homologues thereof (to date, seven TRP homologues and splice variants have been identified in a variety of organisms), the vanilloid receptor subtype I (also known as the capsaicin receptor); the stretch-inhibitable non-selective cation channel (SIC); the olfactory, mechanosensitive channel; the insulin-like growth factor I-regulated calcium channel; the vitamin D-responsive apical, epithelial calcium channel (ECaC); melastatin; and the polycystic kidney disease protein family (see, e.g., Montell and Rubin (1989) *Neuron* 2:1313-1323; Caterina et al. (1997) *Nature* 389: 816-824; Suzuki et al. (1999) *J. Biol. Chem.* 274: 6330-6335; Kiselyov et al. (1998) *Nature* 396: 478-482; Hoenderop et al. (1999) *J. Biol. Chem.* 274: 8375-8378; and Chen et al. (1999) *Nature* 401(6751): 383-6). Each

of these molecules is 700 or more amino acids in length, and shares certain conserved structural features. Predominant among these structural features are six transmembrane domains, with an additional hydrophobic loop present between the fifth and sixth transmembrane domains. It is believed that this loop is integral to the activity of the pore of the channel formed upon membrane insertion (Hardie and Minke (1993) *Trends Neurosci* 16: 371-376). TRP channel proteins also include one or more ankyrin domains. Although found in disparate tissues and organisms, members of the TRP channel protein family all serve to transduce signals by means of calcium entry into cells, particularly pain signals (see, e.g., McClesky and Gold (1999) *Annu. Rev. Physiol.* 61: 835-856), light signals (Hardie and Minke, supra), or olfactory signals (Colbert et al. (1997) *J. Neurosci* 17(21): 8259-8269). Thus, this family of molecules may play important roles in sensory signal transduction in general.

[0186] Vanilloid receptors (VRs) are cation channels that are structurally related to members of the TRP family of ion channels. VRs share several physical characteristics including an N-terminal cytoplasmic domain which contains three ankyrin repeats, six transmembrane domains, a pore-loop region located between transmembrane domains 5 and 6, and several kinase consensus sequences. These receptors have been proposed to mediate the entry of extracellular calcium into cells in response to the depletion of intracellular calcium stores. VRs are expressed in nociceptive neurons, as well as other cells types, and are activated by a variety of stimuli including noxious heat and protons. A well-known agonist of VR1 is capsaicin, which induces pain behavior in humans and rodents. VR1 knockout mice have been shown to be impaired in their detection of painful heat, to exhibit no vanilloid-evoked pain behavior, and to show little thermal hypersensitivity after inflammation (Szallasi and Blumberg (1999) Pharmacol. Rev. 51:159-211).

[0187] The human TLCC-4 or 48000 sequence (SEQ ID NO:21), which is approximately 4586 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 2370 nucleotides, not including the termination codon (nucleotides 146-2515 of SEQ ID NO:21; 1-2370 of SEQ ID NO:23). The coding sequence encodes a 790 amino acid protein (SEQ ID NO:22).

[0188] The human TLCC-5 or 52920 sequence (SEQ ID NO:24), which is approximately 3042 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 3039 nucleotides, not including the termination codon (nucleotides 1-3039 of SEQ ID NO:24; 1-3039 SEQ ID NO:26). The coding sequence encodes a 1013 amino acid protein (SEQ ID NO:25).

[0189] The human TLCC-5 amino acid sequence was aligned with the amino acid sequence of transient receptor potential polypeptide 7 (TRP7) and melastatin from *Homo sapiens* using the CLUSTAL W (1.74) multiple sequence alignment program. That alignment shows the homology of the proteins.

[0190] A search was performed against the HMM database in PFAM resulting in the identification of four ankyrin repeat domains at about residues 167-202 (score=1.6), 214-246 (score=30.6), 261-294 (score=27.9), and 340-372 (score=18.6), and an ion transport protein domain at about residues 510-677 (score=34.5) in the amino acid sequence of human TLCC-4 (SEQ ID NO:22).

[0191] A search was also performed against the MEMSAT database resulting in the identification of six transmembrane domains in the amino acid sequence of human TLCC-4 at about residues 440-461, 488-508, 520-540, 547-565, 590-609, and 652-676 of SEQ ID NO:22.

[0192] A search was further performed against the HMM database resulting in the identification of two transient receptor domains at about residues 720-778 (score=21.7) and 820-876 (score=1.5), in the amino acid sequence of human TLCC-5 (SEQ ID NO:25).

[0193] A search was also performed against the MEMSAT database resulting in the identification of two transmembrane domains in the amino acid sequence of human TLCC-5 at about residues 786-803 and 826-848 of SEQ ID NO:25.

[0194] A search in the Prosite database further resulted in the identification of eight protein kinase C phosphorylation sites in the amino acid sequence of human TLCC-4 (SEQ ID NO:22) at about residues 37-39, 167-169, 290-292, 335-337, 374-376, 476-478, 498-500, and 688-690; two N-glycosylation sites in the amino acid sequence of human TLCC-4 (SEQ ID NO:22) at about residues 452-455 and 683-686; a cAMP- and cGMP-dependent protein kinase phosphorylation site in the amino acid sequence of human TLCC-4 (SEQ ID NO:22) at about residues 375-378; fourteen casein kinase II phosphorylation sites in the amino acid sequence of human TLCC-4 (SEQ ID NO:22) at about residues 88-91, 163-166, 290-293, 305-308, 312-315, 388-391, 393-396, 397-400, 402-405, 411-414, 498-501, 607-610, 624-627, and 699-702; three tyrosine kinase phosphorylation sites in the amino acid sequence of human TLCC-4 (SEQ ID NO:22) at about residues 253-260, 375-382, and 614-622; two N-myristoylation sites in the amino acid sequence of human TLCC-4 (SEQ ID NO:22) at about residues 238-243 and 602-607; an amidation site in the amino acid sequence of human TLCC-4 (SEQ ID NO:22) at about residues 12-15; and a leucine zipper site in the amino acid sequence of human TLCC-4 (SEQ ID NO:22) at about residues 584-605.

[0195] A search performed in the Prosite database further resulted in the identification of thirteen protein kinase C phosphorylation sites in the amino acid sequence of human TLCC-5 (SEQ ID NO:25) at about residues 21-23, 28-30, 39-41, 105-107, 240-242, 305-307, 331-333, 338-340, 711-713, 802-804, 901-903, 972-974, and 1001-1003; twelve casein kinase II phosphorylation sites in the amino acid sequence of human TLCC-5 (SEQ ID NO:25) at about residues 54-57, 143-146, 223-226, 240-243, 308-311, 360-363, 436-439, 487-490, 576-579, 725-728, 977-980, and 982-985; and three tyrosine kinase phosphorylation sites in the amino acid sequence of human TLCC-5 (SEQ ID NO:25) at about residues 49-55, 247-254, and 307-314.

[0196] Further domain motifs were identified by using the amino acid sequence of TLCC-4 (SEQ ID NO:22) to search through the ProDom database. Numerous matches against protein domains described as "receptor vanilloid channel activated receptor-related receptor-like type OTRPC4", "channel vanilloid receptor activated receptor-related receptor-like OTRPC4 2B ion", "repeat ankyrin kinase nuclear factor channel", "ankyrin repeat kinase domain UNC-44 alternative glycoprotein EGF-like", "ankyrin", "channel osmotically receptor-related vanilloid cation", "receptor vanilloid channel activated receptor-related receptor-like calcium type", "calcium epithelial channel transporter homolog CAT2", "channel protein receptor calcium transient potential transmembrane ion transport", and "receptor vanilloid channel activated osmotically", and the like were identified.

[0197] Further domain motifs were identified by using the amino acid sequence of TLCC-5 (SEQ ID NO:25) to search through the ProDom database. Numerous matches against protein domains described as "channel protein calcium entry capacitative ionic transmembrane ion transport transient" and the like were identified.

[0198] As used herein, an "ion channel" includes a protein or polypeptide which is involved in receiving, conducting, and transmitting signals in an electrically excitable cell, e.g., a neuronal or muscle cell. Ion channels include calcium channels, potassium channels, and sodium channels. As used herein, a "calcium channel" includes a protein or polypeptide which is involved in receiving, conducting, and transmitting calcium ion-based signals in an electrically excitable cell. Calcium channels are calcium ion selective, and can determine membrane excitability (the ability of, for example, a neuronal cell to respond to a stimulus and to convert it into a sensory impulse). Calcium channels can also influence the resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation. Calcium channels are typically expressed in electrically excitable cells, e.g., neuronal cells, and may form heteromultimeric structures (e.g., composed of more than one type of subunit). Calcium channels may also be found in non-excitable cells (e.g., adipose cells or liver cells), where they may play a role in, e.g., signal transduction. Examples of calcium channels include the low-voltage-gated channels and the high-voltage-gated channels. Calcium channels are described in, for example, Davila et al. (1999) *Annals New York Academy of Sciences* 868:102-17 and McEnery, M. W. et al. (1998) *J. Bioenergetics and Biomembranes* 30(4): 409-418, the contents of which are incorporated herein by reference. As the TLCC-4 and TLCC-5 molecules of the present invention are calcium channels modulating ion channel mediated activities (e.g., calcium channel- and/or vanilloid receptor-mediated activities), they may be useful for developing novel diagnostic and therapeutic agents for ion channel associated disorders (e.g., calcium channel and/or vanilloid receptor associated disorders).

[0199] As used herein, an "ion channel associated disorder" includes a disorder, disease or condition which is characterized by a misregulation of an ion channel mediated activity. For example, a "calcium channel associated disorder" includes a disorder, disease or condition which is characterized by a misregulation of a calcium channel mediated activity. Ion channel associated disorders, e.g., calcium channel associated disorders, include CNS disorders, such as cognitive and neurodegenerative disorders.

[0200] Ion channel associated disorders, e.g., vanilloid receptor associated disorders also include pain disorders. As used herein, the term "pain disorders" includes those disorders, diseases or conditions that affect pain signaling mechanisms.

[0201] Thus, the TLCC-4 or TLCC-5 molecules by participating in pain signaling mechanisms, may modulate pain elicitation and act as targets for developing novel diagnostic targets and therapeutic agents to control pain.

[0202] Ion channel associated disorders, e.g., calcium channel and/or vanilloid receptor disorders, also include cellular proliferation, growth, differentiation, or migration disorders.

[0203] As used herein, an "ion channel mediated activity" includes an activity which involves an ion channel, e.g., an ion channel and/or a vanilloid receptor, in a neuronal cell, a muscular cell, a skin cell or a liver cell, associated with receiving, conducting, and transmitting signals. Ion channel mediated activities (e.g., calcium channel and/or vanilloid receptor mediated activities) include release of neurotransmitters or second messenger molecules (e.g., dopamine or norepinephrine), from cells, e.g., neuronal cells; modulation of resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation; participation in signal transduction pathways, and modulation of processes such as integration of sub-threshold synaptic responses and the conductance of back-propagating action potentials in, for example, neuronal cells (e.g., changes in those action potentials resulting in a morphological or differentiative response in the cell).

[0204] The term "family" when referring to the polypeptide and nucleic acid molecules of the invention is intended to mean two or more polypeptides or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein. For example, the family of TLCC-4 and TLCC-5 polypeptides comprise at least one "transmembrane domain." As used herein, the term "transmembrane domain" includes an amino acid sequence of about 20-45 amino acid residues in length which spans the plasma membrane. More preferably, a transmembrane domain includes about at least 20, 25, 30, 35, 40, or 45 amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an alpha-helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, e.g., leucines, isoleucines, alanines, valines, phenylalanines, prolines or methionines. Transmembrane domains are described in, for example, Zagotta W. N. et al, (1996) *Annual Rev. Neurosci.* 19: 235-263, the contents of which are incorporated herein by reference. Amino acid residues 440-461, 488-508, 520-540, 547-565, 590-609, and 652-676 of the human TLCC-4 polypeptide (SEQ ID NO:22) comprise transmembrane domains. Amino acid residues 786-803 and 826-848 of the human TLCC-5 polypeptide (SEQ ID NO:25) comprise transmembrane domains. Accordingly, TLCC-4 and/or TLCC-5 polypeptides having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with a transmembrane domain of human TLCC-4 and/or TLCC-5 are within the scope of the invention.

[0205] In another embodiment, a TLCC-4 molecule of the present invention is identified based on the presence of at least one "pore domain" between the fifth and sixth transmembrane domains. As used herein, the term "pore domain" includes an overall hydrophobic amino acid sequence which is located between two transmembrane domains of a calcium channel protein, preferably transmembrane domains 5 and 6, and which is believed to be a major determinant of ion selectivity and channel activity in calcium channels. Pore domains are described in, for example, Vannier et al. (1998) *J. Biol. Chem.* 273: 8675-8679 and Phillips, A. M. et al. (1992) *Neuron* 8, 631-642, the contents of which are incorporated herein by reference. TLCC-4 molecules having at least one pore domain are within the scope of the invention. Amino acid residues 620-640 of the human TLCC-4 sequence (SEQ ID NO:22) comprise a pore domain.

[0206] In another embodiment, a TLCC-4 molecule of the present invention is identified based on the presence of at least

one "ankyrin repeat domain." As used herein, the term "ankyrin repeat domain" includes an amino acid sequence of about 10-110 amino acid residues which serves as an ankyrin repeat. Preferably, an ankyrin repeat domain includes at least about 30 amino acid residues. To identify the presence of an ion transport domain in a TLCC-4 protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein may be searched against a database of known protein domains (e.g., the HMM database). The ankyrin repeat domain (HMM) has been assigned the PFAM Accession PF00023. A search was performed against the HMM database resulting in the identification of ankyrin repeat domains in the amino acid sequence of human TLCC-4 at about residues 167-202, 214-246, 261-294, and 340-372 of SEQ ID NO:22.

[0207] In another embodiment, a TLCC-4 molecule of the present invention is identified based on the presence of at least one "ion transport protein domain." As used herein, the term "ion transport protein domain" includes a protein domain having an amino acid sequence of about 100-200 amino acid residues which serves to transportions. Preferably, an ion transport protein domain includes at least about 160 amino acid residues. To identify the presence of an ion transport protein domain in a TLCC-4 protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein may be searched against a database of known protein domains (e.g., the HMM database). The ion transport domain (HMM) has been assigned the PFAM Accession PF00520. A search was performed against the HMM database resulting in the identification of an ion transport protein domain in the amino acid sequence of human TLCC-4 at about residues 510-677 of SEQ ID NO:22.

[0208] In another embodiment, a TLCC-5 molecule of the present invention is identified based on the presence of at least one "transient receptor domain." As used herein, the term "transient receptor domain" includes a protein domain having an amino acid sequence of about 100-200 amino acid residues which is found in transient receptor potential (Trp) proteins and related ion channel proteins. Preferably, a transient receptor domain includes at least about 56-58 amino acid residues. To identify the presence of a transient receptor domain in a TLCC-5 protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein may be searched against a database of known protein domains (e.g., the HMM database). The transient receptor domain (HMM) has been assigned the PFAM Accession PF02164. A search was performed against the HMM database resulting in the identification of transient receptor domains in the amino acid sequence of human TLCC-5 at about residues 720-778 and 820-876 of SEQ ID NO:25.

[0209] A description of the Pfam database can be found in Sonhammer et al. (1997) *Proteins* 28:405-420 and a detailed description of HMMs can be found, for example, in Gribskov et al. (1990) *Meth. Enzymol.* 183:146-159; Gribskov et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:4355-4358; Krogh et al. (1994) *J. Mol. Biol.* 235:1501-1531; and Stultz et al. (1993) *Protein Sci.* 2:305-314, the contents of which are incorporated herein by reference.

[0210] In a preferred embodiment, the TLCC-4 or TLCC-5 molecules of the invention include at least one transmembrane domain, at least one ankyrin repeat domain, at least one pore domain, at least one transient receptor domain, and/or at least one ion transport protein domain.

[0211] In a preferred embodiment, a TLCC-4 or TLCC-5 polypeptide includes at least one or more of the following domains: an ankyrin repeat domain, and/or a transmembrane domain, and/or a pore domain, and/or a transient receptor domain, and/or an ion transport protein domain, and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homologous or identical to the amino acid sequence of SEQ ID NO:22 or 25. In yet another preferred embodiment, a TLCC-4 or TLCC-5 polypeptide includes at least one or more of the following domains: an ankyrin repeat domain, and/or a transmembrane domain, and/or a pore domain, and/or a transient receptor domain, and/or an ion transport protein domain, and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a complement of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:21, 23, 24, or 26. In another preferred embodiment, a TLCC-4 or TLCC-5 polypeptide includes at least one or more of the following domains: an ankyrin repeat domain, and/or a transmembrane domain, and/or a pore domain, and/or a transient receptor domain, and/or an ion transport protein domain, and has a TLCC-4 or TLCC-5 activity.

[0212] As used interchangeably herein, a "TLCC-4 or TLCC-5 activity", "biological activity of TLCC-4 or TLCC-5" or "functional activity of TLCC-4 or TLCC-5", refers to an activity exerted by a TLCC-4 or TLCC-5 polypeptide or nucleic acid molecule on a TLCC-4 or TLCC-5 responsive cell or tissue, or on a TLCC-4 or TLCC-5 polypeptide substrate, as determined in vivo, or in vitro, according to standard techniques. In one embodiment, a TLCC-4 or TLCC-5 activity is a direct activity, such as an association with a TLCC-4-target molecule or TLCC-5-target molecule. As used herein, a "target molecule" or "binding partner" is a molecule with which a TLCC-4 or TLCC-5 polypeptide binds or interacts in nature, such that TLCC-4-mediated or TLCC-5-mediated function is achieved. A TLCC-4 or TLCC-5 target molecule can be a non-TLCC-4 or non-TLCC-5 molecule or a TLCC-4 or TLCC-5 polypeptide or polypeptide of the present invention. In an exemplary embodiment, a TLCC-4 or TLCC-5 target molecule is a TLCC-4 or TLCC-5 ligand, e.g., a calcium channel ligand such as calcium. Alternatively, a TLCC-4 or TLCC-5 activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the TLCC-4 or TLCC-5 polypeptide with a TLCC-4 or TLCC-5 ligand. The biological activities of TLCC-4 or TLCC-5 are described herein. For example, the TLCC-4 or TLCC-5 polypeptides of the present invention can have one or more of the following activities: (1) modulate membrane excitability, (2) influence the resting potential of membranes, (3) modulate wave forms and frequencies of action potentials, (4) modulate thresholds of excitation, (5) modulate neurite outgrowth and synaptogenesis, (6) modulate signal transduction, and (7) participate in nociception.

[0213] Accordingly, another embodiment of the invention features isolated TLCC-4 or TLCC-5 polypeptides and polypeptides having a TLCC-4 or TLCC-5 activity. Preferred polypeptides are TLCC-4 or TLCC-5 polypeptides having at least one or more of the following domains: an ankyrin repeat domain, and/or a transmembrane domain, and/or a pore domain, and/or a transient receptor domain, and/or an ion transport protein domain and, preferably, a TLCC-4 or TLCC-5 activity.

[0214] Additional preferred polypeptides have one or more of the following domains: an ankyrin repeat domain, a transmembrane domain, a pore domain, a transient receptor domain, and/or an ion transport protein domain, and are, preferably, encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:21, 23, 24, or 26.

Tissue Distribution of TLCC-4 mRNA

[0215] The tissue distribution of TLCC-4 mRNA was determined by RT-PCR using various cDNA libraries and a human TLCC-4-specific probe. From this analysis it was determined that TLCC-4 mRNA was expressed predominantly in the hypothalamus and skin. TLCC-4 mRNA was found in moderate levels in adipose and testes, and in low levels in skeletal muscle and brain.

[0216] In situ hybridization experiments were performed using a human TLCC-4-specific probe indicating TLCC-4 expression in monkey brain (cortex, thalamus, caudate, and hippocampus), spinal cord, DRG and SRG neurons, and in hair follicles. In situ hybridization with rat pain models indicated that TLCC-4 mRNA was down-regulated after chronic constriction injury, which causes persistent, spontaneous firing of neurons and results in pain. TLCC-4 mRNA was also down-regulated after treatment with clofibric acid, a selective muscle toxin which produces muscle pain and inflammation.

Gene Expression of TLCC-4

[0217] The tissue distribution of human TLCC-4 mRNA was also determined using the TaqMan procedure on a variety of cells and tissues.

[0218] Strong expression of TLCC-4 was detected in human brain (hypothalamus) and skin tissues. In addition, TLCC-4 expression was detected at moderate levels in adipose and testis tissues, and at low levels in the fetal heart, skeletal muscle, brain, and colon tissues. Pain human panel phase I and MP Phase 1.3.3 libraries were also analyzed and it was determined that TLCC-4 was expressed at high levels in the brain, cortex, and testis, at moderate levels in the spinal cord, dorsal root ganglion (DRG), and the hypothalamus, and at low levels in the skin, placenta, small intestine, ovary, prostate epithelial cells, liver, skin (decubitus), colon tumor cells, and breast tumor cells. Monkey libraries were also analyzed indicating that TLCC-4 was expressed at high levels in the monkey cortex and hairy skin, and at low levels in the monkey spinal cord. Metabolic libraries were also analyzed demonstrating that TLCC-4 was expressed at high levels in adipose and brain tissues, and at low levels in differentiated adipocytes and pre-adipocytes, as well as in the hypothalamus, colon, small intestine, skeletal muscle, and liver tissues.

Regulation of Calcium Influx Through TLCC-4

[0219] The regulation of calcium influx though TLCC-4 in 911 cells was determined by Fluorometric Imaging Plate Reader experiments (FLIPR) (Molecular Devices Corp., Sunnyvale, Calif.).

[0220] The FLIPR is a screening tool for cell-based fluorescent assays which allows the simultaneous stimulation and measurement of separate cell populations in a high throughput format. Therefore, using this system, it is possible to quantify transient signals, such as the release of intracellular calcium, from cell populations, in parallel and in real time. The FLIPR contains chambers in which to hold the test plate and plates containing antagonists or agonists to be added to the test plate. The FLIPR utilizes an argon laser that provides discrete spectral lines spaced from approximately 350 to 530 nm. For use with fluorescent Ca^{2+} dyes, the 88-nm line of the laser is employed. The laser simultaneously illuminates the wells in a test plate. The image of each well in the plate is captured by a cooled charge coupled device (CCD) camera, which updates images once per second, if required, for the measurement of rapid calcium responses. Because both excitation and emission are read via the bottom of the plate, black-walled, transparent bottomed 96-well plates are used. Data captured by the CCD camera is converted to digital data and then transferred to a computer.

[0221] Briefly, a calcium indicator (e.g., fluo-3/AM or Calcium Green-1/AM) was transferred to the culture medium. Because the FLIPR collects fluorescence from the bottom of the well. Suspension cells require centrifugation to the base of the well following dye loading. Viable 911 cells were resuspended in loading medium and incubated for one hour. The cells were then centrifuged and resuspended with wash buffer. The cell suspension containing the dye was then aliquoted into each well of the black-walled, transparent bottomed 96-well plate and the plate was centrifuged. The FLIPR assay was then carried out and the results analyzed. (If adherent cells are used, they may be plated at an appropriate density in the 96-well plates and cultured overnight. Dye may then be loaded and incubated).

[0222] Results show a constitutive calcium influx through TLCC-4 in 911 cells that were incubated with NMDG/0 Ca^{2+} and stimulated afterwards with 5 mM Ca^{2+}.

Human 5433

[0223] Non-voltage gated cation currents, which are activated following stimulation of phospholipase C (PLC), appear to be major modes for Ca^{++} and Na^{+} entry in cells. The TRP channel family may mediate some of these conductances since their expression in vitro leads to PLC-dependent calcium influx. Members of this family have been proposed to mediate the entry of extracellular calcium into cells in response to the depletion of intracellular calcium stores. The influx of Ca^{++} and Na^{+} is essential in the nervous system for propagation of action potential, synaptic transmission etc. Furthermore, one member of the channel family, TRPC3 or 5433, has been shown to contribute to a PLC-dependent calcium influx induced by brain-derived neurotrophic factor (BDNF), a known mediator of neuropathic pain. Thus, this family of molecules may play important roles in sensory signal transduction in general.

[0224] The human 5433 sequence (SEQ ID NO:27), which is approximately 3448 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 2544 nucleotides (nucleotides 425-2968 of SEQ ID NO:27; 1-2544 of SEQ ID NO:29), not including the termination codon. The coding sequence encodes a 848 amino acid protein (SEQ ID NO:28).

[0225] Human 5433 contains the following regions or other structural features: one predicted calcium channel domain (ion transport protein domain, PFAM Accession Number PF00520) located at about amino acid residues 436 to 670 of SEQ ID NO:28; two predicted ank repeat domains (PFAM Accession Number PF00023) located at about amino acids 73 to 105, and 159 to 191 of SEQ ID NO:28; six predicted transmembrane segments located at about amino acids 386 to 402, 434 to 450, 474 to 492, 541 to 557, 580 to 603, and 646

to 670 of SEQ ID NO:28; one predicted N-terminal cytoplasmic domain located at about amino acids 1 to 385 of SEQ ID NO:28; one predicted C-terminal cytoplasmic domain located at about amino acids 671 to 848 of SEQ ID NO:28; two predicted cytoplasmic loops located at about amino acids 451 to 473, and 558 to 579 of SEQ ID NO:28; three predicted non-cytoplasmic loops located at about amino acids 403 to 433, 493 to 540, and 604 to 645 of SEQ ID NO:28; six predicted N-glycosylation sites (PS00001) located at about amino acids 337 to 340, 403 to 406, 416 to 419, 560 to 563, 655 to 658, and 671 to 674 of SEQ ID NO:28; three predicted cAMP- and cGMP-dependent protein kinase phosphorylation sites (PS00004) located at about amino acids 8 to 11, 131 to 134, and 260 to 263 of SEQ ID NO:28; nine predicted Protein Kinase C phosphorylation sites (PS00005) located at about amino acids 6 to 8, 130 to 132, 405 to 407, 485 to 487, 573 to 575, 617 to 619, 712 to 714, 813 to 815, and 837 to 839 of SEQ ID NO:28; eleven predicted Casein Kinase II phosphorylation sites (PS00006) located at about amino acids 37 to 40, 136 to 139, 230 to 233, 237 to 240, 418 to 421, 435 to 438, 573 to 576, 612 to 615, 673 to 676, 681 to 684, and 695 to 698 of SEQ ID NO:28; two predicted Tyrosine kinase phosphorylation sites (PS00007) located at about amino acids 42 to 49, and 641 to 648 of SEQ ID NO:28; four predicted N-myristylation sites (PS00008) located at about amino acids 33 to 38, 447 to 452, 604 to 609, and 741 to 746 of SEQ ID NO:28; and one predicted Amidation site (PS00009) located at about amino acids 16 to 19 of SEQ ID NO:28.

[0226] For general information regarding PFAM identifiers, PS prefix and PF prefix domain identification numbers, refer to Sonnhammer et al. (1997) *Protein* 28:405-420.

TABLE 6

Summary of transmembrane domains.

Transmembrane segment located in SEQ ID NO: 28	Cytoplasmic sequence located in SEQ ID NO: 28	Extracellular sequence located in SEQ ID NO: 28
about 386 to 402	about 1 to 385 (N-terminal)	
about 434 to 450		about 403 to 433
about 474 to 492	about 451 to 473	
about 541 to 557		about 493 to 540
about 580 to 603	about 558 to 579	
about 646 to 670		about 604 to 645
	about 671 to 848 (C-terminal)	

[0227] A hydropathy plot of human 5433 was performed. Polypeptides of the invention include fragments which include: all or part of a hydrophobic sequence, e.g., the sequence from about amino acid 265 to 275, from about 345 to 380, and from about 702 to 720 of SEQ ID NO:28; all or part of a hydrophilic sequence, e.g., the sequence of from about amino acid 194 to 220, from about 245 to 262, and from about 720 to 741 of SEQ ID NO:28; a sequence which includes a Cys, or a glycosylation site.

[0228] The 5433 protein contains a significant number of structural characteristics in common with members of the calcium channel family. In particular, the 5433 protein shows homology to the transient receptor potential (TRP) proteins. The term "family" when referring to the protein and nucleic acid molecules of the invention means two or more proteins or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein.

[0229] As used herein, a "calcium channel" includes a protein or polypeptide that is involved in receiving, conducting, and transmitting signals in an electrically excitable cell, e.g., a neuronal or muscular cell. Calcium channels are calcium ion selective, and can determine membrane excitability (the ability of, for example, a muscle cell to respond to a stimulus and to convert it into an impulse resulting in a contraction). Calcium channels can also influence the resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation. Calcium channels are typically expressed in electrically excitable cells, e.g., neuronal or muscle cells, and may form heteromultimeric structures (e.g., composed of more than one type of subunit). Calcium channels are described in, for example, Davila et al. (1999) Annals New York Academy of Sciences 868: 102-17 and McEnery, M. W. et al. (1998) *J. Bioenergetics and Biomembranes* 30(4): 409-418, the contents of which are incorporated herein by reference.

[0230] Calcium signaling has been implicated in the regulation of a variety of cellular responses, such as neuronal development and maintenance, and cell growth and differentiation. There are two general methods by which intracellular concentrations of calcium ions may be increased: calcium ions may be brought into the cell from the extracellular milieu through the use of specific channels in the cellular membrane, or calcium ions may be freed from intracellular stores, again being transported by specific membrane channels in the storage organelle. In the situation in which the intracellular stores of calcium have been depleted, a specific type of calcium channel, termed a "capacitative calcium channel" or a "store-operated calcium channel" (SOC), is activated in the plasma membrane to import calcium ions from the extracellular environment to the cytosol (for review, see Putney and McKay (1999) *BioEssays* 21: 38-46). Members of the capacitative calcium channel family include the calcium release-activated calcium current (Hoth and Penner (1992) *Nature* 355: 353-355), calcium release-activated nonselective cation current (Krause et al. (1996) *J. Biol. Chem.* 271: 32523-32528), and the transient receptor potential (TRP) proteins. See, e.g., Putney, J. W., (1986) *Cell Calcium* 7: 1-12; Putney, J. W. (1990) *Cell Calcium* 11: 611-624.

[0231] The TRP channel family is one of the best characterized of the capacitative calcium channel group. These channels include transient receptor potential protein and homologues thereof, the vanilloid receptor subtype I, stretch-inhibitable non-selective cation channel, olfactory, mechanosensitive channel, insulin-like growth factor I-regulated calcium channel, and vitamin D-responsive apical, epithelial calcium channel (ECaC) (see, e.g., Montell and Rubin (1989) *Neuron* 2:1313-1323; Caterina et al. (1997) *Nature* 389: 816-824; Suzuki et al. (1999) *J. Biol. Chem.* 274: 6330-6335; Kiselyov et al. (1998) *Nature* 396: 478-482; and Hoenderop et al. (1999) *J. Biol. Chem.* 274: 8375-8378). Each of these molecules is at least 700 amino acids, and shares certain conserved structural features. Predominant among these structural features are six transmembrane domains, with an additional hydrophobic loop present between the fifth and sixth transmembrane domains. It is believed that this loop is integral to the activity of the pore of the channel formed upon membrane insertion (Hardie and Minke (1993) *Trends Neurosci* 16: 371-376). TRP channel proteins also include one or more ankyrin domains and frequently display a proline-rich region at the N-terminus. See, e.g., McClesky and Gold

(1999) *Annu. Rev. Physiol.* 61: 835-856; and Colbert et al. (1997) *J. Neurosci* 17(21): 8259-8269.

[0232] Non-voltage gated cation currents, which are activated following stimulation of phospholipase C (PLC), appear to be major modes for Ca^{++} and Na^{+} entry in cells. The TRP channel family may mediate some of these conductances since their expression in vitro leads to PLC-dependent calcium influx. Members of this family have been proposed to mediate the entry of extracellular calcium into cells in response to the depletion of intracellular calcium stores. The influx of Ca^{++} and Na^{+} is essential in the nervous system for propagation of action potential, synaptic transmission etc. Furthermore, one member of the channel family, TRPC3, has been shown to contribute to a PLC-dependent calcium influx induced by brain-derived neurotrophic factor (BDNF), a known mediator of neuropathic pain. Thus, this family of molecules may play important roles in sensory signal transduction in general.

[0233] The 5433 protein shows homology to transient receptor potential (TRP) proteins. The term "transient receptor potential" protein refers to a membrane-spanning, glycoprotein cation (calcium or sodium) channels. A limited sequence similarity to voltage-gated calcium channel al subunits lead to the prediction of six transmembrane segments flanked by intracellular N and C termini and a putative pore region between the transmembrane segments 5 and 6. Generally, the first hydrophobic region rather than being a transmembrane segment is intracellular and available for protein-protein interactions. See, e.g., Vannier B. et al. (1998) *J. Biol. Chem.* 273: 8675-8679. Preferably, the TRP protein includes six endogenous glycosylation sites (Zhu X. et al. (1996) *Cell* 85: 661-671).

[0234] The 5433 polypeptides contain structural features similar to calcium channel protein family members, in particular, to the TRP proteins. For example, the 5433 polypeptide has seven predicted hydrophobic regions present at about amino acids 350 to 369, 386 to 402, 434 to 450, 474 to 492, 541 to 557, 580 to 603, and 646 to 670 of SEQ ID NO:28, and the first region rather than being a transmembrane segment is intracellular domain. Accordingly, the 5433 polypeptide has six transmembrane segments flanked by intracellular N and C termini. The 5433 polypeptide also has six predicted N-glycosylation sites (PS00001) located at about amino acids 337 to 340, 403 to 406, 416 to 419, 560 to 563, 655 to 658, and 671 to 674 of SEQ ID NO:28.

[0235] Non-voltage gated cation currents, which are activated following stimulation of phospholipase C, appear to be major modes for calcium and sodium entry in cells. TRP family members are believed to mediate some of these conductances since their expression in vitro leads to phospholipase C (PLC)-dependent calcium influx. Members of this family have been proposed to mediate entry of extracellular calcium into cells in response to the depletion of intracellular stores. The influx of calcium and sodium is essential in the nervous system for the propagation of action potentials, synaptic transmission, etc. Furthermore, TRPs have been shown to contribute to PLC-dependent calcium influx induced by BDNF, a known mediator of neuropathic pain. Accordingly, 5433-activity may be involved in neurological processes, including PLC-mediated conductances associated with the propagation of action potentials, synaptic transmission, nociceptive responses, and neuropathic pain.

[0236] A 5433 polypeptide can include an "ion channel domain" or a "calcium channel domain" (or "ion transport protein domain"), or regions homologous with a "calcium channel domain."

[0237] As used herein, the term "calcium channel domain" includes an amino acid sequence of about 150 to 450 amino acid residues in length and having a bit score for the alignment of the sequence to the calcium channel domain profile (PFAM HMM) of at least 50. Preferably, a calcium channel domain includes at least about 200 to 300 amino acids, more preferably about 210 to 250 amino acid residues, or about 234 amino acids and has a bit score for the alignment of the sequence to the calcium channel domain (HMM) of at least 60, preferable 70, 75 or greater. The calcium channel domain (HMM) has been assigned the PFAM Accession Number PF00520. The calcium channel domain (amino acids 436 to 670 of SEQ ID NO:28) of human 5433 aligns with a consensus amino acid sequence (SEQ ID NO:30) derived from a hidden Markov model.

[0238] In a preferred embodiment 5433 polypeptide or protein has a "calcium channel domain" or a region which includes at least about 150 to 450, more preferably about 200 to 300 amino acids, more preferably about 210 to 250 amino acid residues and has at least about 70% 80% 90% 95%, 99%, or 100% homology with a "calcium channel domain," e.g., the calcium channel domain of human 5433 (e.g., residues 436 to 670 of SEQ ID NO:28).

[0239] To identify the presence of a "calcium channel" domain in a 5433 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against the Pfam database of HMMs (e.g., the Pfam database, release 2.1) using the default parameters. For example, the hmmsf program, which is available as part of the HMMER package of search programs, is a family specific default program for MILPAT0063 and a score of 15 is the default threshold score for determining a hit. Alternatively, the threshold score for determining a hit can be lowered (e.g., to 8 bits). A description of the Pfam database can be found in Sonhammer et al. (1997) *Proteins* 28(3): 405-420 and a detailed description of HMMs can be found, for example, in Gribskov et al. (1990) *Meth. Enzymol.* 183: 146-159; Gribskov et al. (1987) *Proc. Natl. Acad. Sci. USA* 84: 4355-4358; Krogh et al. (1994) *J. Mol. Biol.* 235: 1501-1531; and Stultz et al. (1993) *Protein Sci.* 2: 305-314, the contents of which are incorporated herein by reference. A search was performed against the HMM database resulting in the identification of a "calcium channel" domain in the amino acid sequence of human 5433 at about residues 436 to 670 of SEQ ID NO:28.

[0240] A 5433 molecule can further include at least one, or two ank repeat domains. An ank repeat domain is characterized by a common fold, of about 30 amino acids, characterized by a helix-beta-turn-helix core. See, for example, Kalus W. et al. (1997) *FEBS Lett* 401(2-3): 127-32.

[0241] As used herein, the term "ank repeat domain" (or "ankyrin domain") includes an amino acid sequence of about 10 to 50 amino acid residues in length and having a bit score for the alignment of the sequence to the ank repeat domain (HMM) of at least 5. Preferably, an ank domain includes at least about 20 to 40 amino acids, more preferably about 30 to 35 amino acids, or about 32 amino acids, and has a bit score for the alignment of the sequence to the ank repeat domain (HMM) of at least 10, preferably 14, or more preferably 15 or greater. The ank repeat domain (HMM) has been assigned the

PFAM Accession Number PF00023. The ank repeat domain (amino acids 73 to 105, or 159 to 191 of SEQ ID NO:28) of human 5433 aligns with a consensus amino acid sequence (SEQ ID NO:31) derived from a hidden Markov model.

[0242] In a preferred embodiment 5433 polypeptide or protein has an "ank repeat domain" or a region which includes at least about 10 to 50 more preferably about 20 to 40, or about 32 amino acid residues and has at least about 70% 80% 90% 95%, 99%, or 100% homology with an "ank repeat domain," e.g., the ank repeat domain of human 5433 (e.g., amino acids 73 to 105, or 159 to 191 of SEQ ID NO:28).

[0243] To identify the presence of an "ank repeat" domain in a 5433 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against a database of HMMs as described above. A search was performed against the HMM database resulting in the identification of an "ank repeat domain" domain in the amino acid sequence of human 5433 at about residues acids 73 to 105, or 159 to 191 of SEQ ID NO:28.

[0244] A 5433 protein further includes a predicted N-terminal cytoplasmic domain located at about amino acids 1-385 of SEQ ID NO:28. As used herein, a "N-terminal cytoplasmic domain" includes an amino acid sequence having about 1-500, preferably about 1-400, or even more preferably about 1-390 amino acid residues in length and is located inside of a cell or intracellularly. The C-terminal amino acid residue of a "N-terminal cytoplasmic domain" is adjacent to a N-terminal amino acid residue of a transmembrane domain in a 5433 protein. For example, a N-terminal cytoplasmic domain is located at about amino acid residues 1-385 of SEQ ID NO:28.

[0245] In a preferred embodiment 5433 polypeptide or protein has an "N-terminal cytoplasmic domain" or a region which includes at least about 1-600, preferably about 100-400, and even more preferably about 385 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with an "N-terminal cytoplasmic domain," e.g., the N-terminal cytoplasmic domain of human 5433 (e.g., residues 1-385 of SEQ ID NO:28).

[0246] In another embodiment, a 5433 protein includes a "C-terminal cytoplasmic domain," also referred to herein as a C-terminal cytoplasmic tail, in the sequence of the protein. As used herein, a "C-terminal cytoplasmic domain" includes an amino acid sequence having a length of at least about 100, more preferably 150 or more amino acid residues and is located within a cell or within the cytoplasm of a cell. Accordingly, the N-terminal amino acid residue of a "C-terminal cytoplasmic domain" is adjacent to a C-terminal amino acid residue of a transmembrane domain in a naturally-occurring 5433 protein. For example, a C-terminal cytoplasmic domain is found at about amino acid residues 671 to 848 of SEQ ID NO:28.

[0247] In a preferred embodiment, a 5433 polypeptide or protein has a C-terminal cytoplasmic domain or a region which includes at least about 100, more preferably 150 or more amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with an "C-terminal cytoplasmic domain," e.g., the C-terminal cytoplasmic domain of human 5433 (e.g., residues 671 to 848 of SEQ ID NO:28).

[0248] 5433 proteins can further include at least one, two, three, four, five, and preferably six transmembrane domains. As used herein, the term "transmembrane domain" includes an amino acid sequence of about 10 to 45, preferably 12 to 30, and most preferably 15 to 25, amino acid residues in length that spans the plasma membrane. More preferably, a transmembrane domain includes about at least 17, 19, 24, or 25 amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an alpha-helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, e.g., leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains are described in, for example, Zagotta W. N. et al, (1996) *Annual Rev. Neurosci.* 19: 235-263, the contents of which are incorporated herein by reference. Amino acid residues 386 to 402, 434 to 450, 474 to 492, 541 to 557, 580 to 603, and 646 to 670 of SEQ ID NO:28 are transmembrane domains. Accordingly, proteins having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, about 80-90%, or about 90-100% homology with amino acids 386 to 402, 434 to 450, 474 to 492, 541 to 557, 580 to 603, and 646 to 670 of SEQ ID NO:28 are within the scope of the invention.

[0249] In another embodiment, a 5433 protein includes at least one, or two cytoplasmic loop, also referred to herein as a cytoplasmic domain. As used herein, a "cytoplasmic loop" includes an amino acid sequence having a length of at least about 10, preferably about 20, amino acid residues located within a cell or within the cytoplasm of a cell. For example, a cytoplasmic loop is found at about amino acids 451 to 473, or 558 to 579 of SEQ ID NO:28.

[0250] In a preferred embodiment 5433 polypeptide or protein has at least one cytoplasmic loop or a region which includes at least about 10, preferably about 20 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with an "cytoplasmic loop," e.g., at least one cytoplasmic loop of human 5433 (e.g., residues 451 to 473, or 558 to 579 of SEQ ID NO:28).

[0251] In another embodiment, a 5433 protein include at least one, two, or three non-cytoplasmic (extracellular) loop. As defined herein, the term "loop" includes an amino acid sequence that resides outside of a phospholipid membrane, having a length of at least about 20 to 70, and preferably about 30 to 50 amino acid residues, and has an amino acid sequence that connects two transmembrane domains within a protein or polypeptide. Extracellular domains are located outside of the cell. Accordingly, the N-terminal amino acid of a non-cytoplasmic loop is adjacent to a C-terminal amino acid of a transmembrane domain in a 5433 protein, and the C-terminal amino acid of a non-cytoplasmic loop is adjacent to an N-terminal amino acid of a transmembrane domain in a 5433 protein. For example, an "extracellular loop" can be found at about amino acids 403 to 433, 493 to 540, and 604 to 645 of SEQ ID NO:28.

[0252] In a preferred embodiment, a 5433 polypeptide or protein has at least one, two, or three non-cytoplasmic loops or regions which include at least about 20 to 70, and preferably about 30 to 50 amino acid residues and has at least about 70% 80% 90% 95%, 99%, or 100% homology with an "extracellular loop," e.g., at least one non-cytoplasmic loop of human 5433 (e.g., residues 403 to 433, 493 to 540, and 604 to 645 of SEQ ID NO:28).

[0253] Accordingly, in one embodiment of the invention, a 5433 includes at least one, two, three, four, five, preferably six, transmembrane domains, at least one, or two cytoplasmic loops, and/or at least one, two, or three extracellular loops. In

another embodiment, the 5433 further includes an N-terminal and a C-terminal cytoplasmic domains.

[0254] A 5433 family member can include at least one predicted calcium channel domain; and at least one, preferably two predicted ank repeat domains. Furthermore, a 5433 family member can include at least one, two, three, four, five, preferably six predicted N-glycosylation sites (PS00001); at least one, two, preferably three predicted cAMP- and cGMP-dependent protein kinase phosphorylation sites (PS00004); at least one, two, three, four, five, six, seven, eight, preferably nine predicted protein kinase C phosphorylation sites (PS00005); at least one, two, three, four, five, six, seven, eight, nine, ten, and preferably eleven predicted casein kinase II phosphorylation sites (PS00006); at least one, preferably two Tyrosine kinase phosphorylation sites (PS00007); at least one, two, three, and preferably four predicted N-myristylation sites (PS00008); and at least one predicted Amidation site (PS00009).

[0255] As the 5433 polypeptides of the invention may modulate 5433-mediated activities, they may be useful as of for developing novel diagnostic and therapeutic agents for 5433-mediated or related disorders, as described below.

[0256] As used herein, a "5433 activity," "biological activity of 5433," or "functional activity of 5433," refers to an activity exerted by a 5433 protein, polypeptide or nucleic acid molecule on e.g., a 5433-responsive cell or on a 5433 substrate, e.g., a protein substrate, as determined in vivo or in vitro. In one embodiment, a 5433 activity is a direct activity, such as an association with a 5433 target molecule. A "target molecule" "substrate" or "binding partner" is a molecule with which a 5433 protein binds or interacts in nature. A 5433 activity can also be an indirect activity, e.g., a cellular signaling activity mediated by interaction of the 5433 protein with a 5433 binding partner. In an exemplary embodiment, 5433 is controlling neurotransmitter release from neurons.

[0257] Based on the above-described sequence similarities and the tissue distribution described below, the 5433 molecules of the present invention are predicted to have similar biological activities as calcium channel family members. Thus, in accordance with the invention, a 5433 calcium channel or subsequence or variant polypeptide may have one or more domains and, therefore, one or more activities or functions characteristic of a calcium channel family member, including, but not limited to, (1) controlling neurotransmitter release from neurons; (2) regulating nociceptive responses; (3) regulating synaptic transmission; (3) modulating cation, e.g., calcium or sodium, entry into a cell, e.g., a neuronal cell; or (5) modulating pain or inflammation response. Thus, the 5433 molecules can act as novel diagnostic targets and therapeutic agents for controlling ion (e.g., calcium) channel-associated disorders.

[0258] Nociceptive responses in general are heterogeneous processes involving intracellular signaling mediated by phospholipase C, a known activator of TRP receptor channels. TRP receptor channels has been shown to be regulated by brain-derived neurotrophic factor (BDNF), a known mediator of neuropathic pain. Thus, 5433 calcium channel protein or subsequence or variant having calcium channel activity is capable of, e.g., modulating nociceptive responses, and also pain responses.

[0259] Expression of 5433 mRNA is detected in human and rat brain, spinal cord, and dorsal root ganglia (DRG) (Tables 7-11). Expression was low in any other normal human and rat tissues. In situ hybridization experiments with the human probe showed expression in monkey and rat cortex, spinal cord, and DRG neurons. Hence, 5433 is likely a neuro-related calcium channel protein involved in neurological response, e.g., nociceptive and pain responses.

[0260] Animal models indicate a role for the 5433 molecule in pain response. Examples of animal models of pain response that can be tested include, but are not limited to, axotomy, the cutting or severing of an axon; chronic constriction injury (CCI), a model of neuropathic pain which involves ligation of the sciatic nerve in rodents, e.g., rats; or intraplantar Freund's adjuvant injection as a model of arthritic pain. Other animal models of pain response are described in, e.g., *ILAR Journal* (1999) Volume 40, Number 3 (entire issue). Taqman experiments in rat animal models show no regulation in DRGs. However, 5433 is up-regulated in the spinal cord after CCI axotomy, and after CFA intraplantar injection.

[0261] Therefore, 5433 associated disorders can detrimentally affect regulation and modulation of the pain response; and vasoconstriction response and pain therefrom. Examples of 5433 associated disorders in which the 5433 molecules of the invention may be directly or indirectly involved include pain, pain syndromes, and inflammatory disorders, including inflammatory pain. Accordingly, the 5433 molecules can act as novel diagnostic targets and therapeutic agents controlling neurological, e.g., neurodegenerative, disorders and pain disorders.

[0262] Agents that modulate 5433 polypeptide or nucleic acid activity or expression can be used to treat pain elicited by any medical condition. A subject receiving the treatment can be additionally treated with a second agent, e.g., an anti-inflammatory agent, an antibiotic, or a chemotherapeutic agent, to further ameliorate the condition.

[0263] The 5433 molecules can also act as novel diagnostic targets and therapeutic agents for brain or neurological disorders.

[0264] 5433 RNA expression is also detected in a panel of cardiovascular (CV) organ and vessel tissues (Tables 12-13). In Table 12, Taqman with CV organ panel shows expression of 5433 in the fetal heart. In Table 13, Taqman with CV vessel panel shows expression of 5433 blood vessels, e.g., artery, Huvec, aorta, and vein. These data show that 5433 calcium channel protein is capable of modulating cardiovascular-related disorders. A cardiovascular disease or disorder also can include an endothelial cell disorder.

Tissue Distribution of 5433 mRNA by TaqMan Analysis

[0265] Taqman experiments using a panel of human normal and tumor tissues, are described in Table 7. The expression of 5433 mRNA using additional human tissues is depicted in Table 8. Taqman experiments using rat panels are depicted in Tables 9-11.

[0266] In humans, 5433 mRNA was highly expressed in the normal brain cortex, hypothalamus, dorsal root ganglion (DRG), and prostate/testis. Further analysis shows expression of 5433 mRNA to be observed in brain, followed by testis, spinal cord, and DRG.

[0267] In rat, expression of 5433 mRNA was observed in the brain, DRG, spinal cord, SGC, and optic nerve. Expression was very low in other normal human and rat tissues tested. Taqman experiments in rat models, respectively, showed no significant regulation of 5433 mRNA in the DRGs. However, 5433 mRNA is upregulated in the spinal cord after CCI, axotomy and after CFA intraplantar injection. The relative tissue distribution of 5433 mRNA is depicted in tabular form in Tables 9-11.

TABLE 7

Tissue Type	Mean	β 2 Mean	∂∂ Ct	Expression
Artery normal	37.24	22.67	12.78	0
Aorta diseased	40	22.98	15.22	0
Vein normal	40	21.34	16.86	0
Coronary SMC	40	22.23	15.98	0
HUVEC	38.54	22.68	14.07	0
Hemangioma	37.06	21.36	13.89	0
Heart normal	37.23	21.41	14.03	0
Heart CHF	38.94	21.68	15.47	0
Kidney	37.41	21.16	14.46	0
Skeletal Muscle	40	23.43	14.77	0
Adipose normal	40	22.7	15.5	0
Pancreas	40	23.6	14.6	0
primary osteoblasts	40	21.3	16.91	0
Osteoclasts (diff)	40	18.26	19.94	0
Skin normal	40	23.28	14.92	0
Spinal cord normal	39.56	22.29	15.47	0
Brain Cortex normal	30.3	23.42	5.08	29.5643
Brain Hypothalamus normal	34.9	24	9.11	1.816
Nerve	37.1	22.51	12.79	0
DRG (Dorsal Root Ganglion)	32.52	22	8.71	2.3797
Breast normal	40	22.11	16.09	0
Breast tumor	40	21.95	16.25	0
Ovary normal	39.95	21.27	16.87	0
Ovary Tumor	39.67	20.1	17.77	0
Prostate Normal	36.98	20.2	14.98	0
Prostate Tumor	38.45	21.12	15.53	0
Salivary glands	39.95	20.36	17.79	0
Colon normal	40	20.18	18.02	0
Colon Tumor	40	22.68	15.53	0
Lung normal	35.23	19.74	13.69	0
Lung tumor	36.56	21.1	13.66	0
Lung COPD	35.32	19.41	14.11	0
Colon IBD	40	18.72	19.48	0
Liver normal	40	21.3	16.9	0
Liver fibrosis	40	22.15	16.06	0
Spleen normal	39.81	19.96	18.05	0
Tonsil normal	37.44	19.16	16.48	0
Lymph node normal	39.39	21.02	16.57	0
Small intestine normal	35.09	21.15	12.15	0
Macrophages	40	18.15	20.05	0
Synovium	40	21.27	16.93	0
BM-MNC	40	20.42	17.78	0
Activated PBMC	36.02	18.59	15.64	0
Neutrophils	40	18.86	19.34	0
Megakaryocytes	40	19.68	18.52	0
Erythroid	40	22.02	16.18	0
positive control	30.72	21.56	7.36	6.0872

TABLE 8

Tissue Type	5433	β2.803	∂Ct	Expression
Adrenal Gland	40.00	19.75	20.26	0.00
Brain	29.73	21.21	8.52	2.73
Heart	39.69	20.31	19.39	0.00
Kidney	38.68	20.31	18.37	0.00
Liver	39.43	20.46	18.97	0.00
Lung	34.02	18.86	15.17	0.03
Mammary Gland	40.00	21.84	18.17	0.00
Pancreas	40.00	21.75	18.26	0.00
Placenta	35.69	21.23	14.46	0.04
Prostate	36.85	19.06	17.80	0.00
Salivary Gland	38.51	20.69	17.83	0.00
Muscle	40.00	22.91	17.10	0.01
Sm. Intestine	36.65	20.26	16.40	0.01
Spleen	37.69	18.55	19.14	0.00
Stomach	34.63	19.33	15.31	0.02
Teste	32.56	21.64	10.92	0.52
Thymus	40.00	26.20	13.80	0.07
Trachea	40.00	22.37	17.63	0.00
Uterus	35.09	20.23	14.86	0.03

TABLE 8-continued

Tissue Type	5433	β2.803	∂Ct	Expression
Spinal Cord	32.51	19.30	13.21	0.11
DRG	34.33	20.24	14.09	0.06
Skin	39.24	19.96	19.28	0.00

TABLE 9

Tissue	r5433	18S	∂Ct	Expression
Brain	27.14	11.12	16.02	0.015
Spinal Cord	29.76	11.42	18.34	0.003
DRG	28.13	11.51	16.62	0.010
SCG	30.19	11.78	18.42	0.003
Optic Nerve	29.94	10.90	19.05	0.002
Hairy Skin	33.83	12.21	21.62	0.000
Gastro Muscle	32.77	11.37	21.40	0.000
Heart	31.47	11.00	20.47	0.001
Kidney	34.01	12.53	21.48	0.000
Liver	39.22	12.09	27.13	0.000
Lung	32.37	12.28	20.09	0.001
Spleen	33.76	11.98	21.78	0.000
Aorta	31.94	11.61	20.34	0.001
Adrenal Gland	32.90	11.40	21.50	0.000
Salivary Gland	34.50	11.21	23.29	0.000
Thyroid	31.62	11.33	20.29	0.001
Prostate	38.23	11.35	26.88	0.000
Thymus	34.43	11.17	23.26	0.000
Trachea	32.69	11.92	20.77	0.001
Esophagus	32.85	11.52	21.33	0.000
Duodenum	33.95	12.37	21.59	0.000
Diaphragm	33.21	12.38	20.83	0.001
Colon	34.05	12.49	21.56	0.000

TABLE 10

Tissue	r5433	18S	∂Ct	Expression
Naïve DRG	28.38	11.75	16.64	0.010
I DRG CCI 3	28.06	11.59	16.48	0.011
I DRG CCI 7	28.13	11.75	16.38	0.012
I DRG CCI 10	28.33	11.77	16.56	0.010
I DRG CCI 28	27.97	11.95	16.02	0.015
Naïve DRG	28.38	11.68	16.71	0.009
I DRG CFA 1	28.91	12.01	16.90	0.008
I DRG CFA 3	28.41	11.98	16.44	0.011
I DRG CFA 7	28.01	11.94	16.08	0.014
I DRG CFA 10	28.20	11.96	16.24	0.013
I DRG CFA 14	27.99	11.55	16.44	0.011
I DRG CFA 28	28.07	11.61	16.46	0.011
Naïve DRG	28.18	11.83	16.35	0.012
I DRG AXT 1	28.04	11.56	16.48	0.011
I DRG AXT 3	28.27	11.74	16.54	0.011
I DRG AXT 7	28.44	11.59	16.86	0.008
I DRG AXT 14	28.23	11.74	16.49	0.011

TABLE 11

Tissue	r5433	18S	∂Ct	Expression
Naïve SC	27.21	11.60	15.62	0.12
I SC CCI 3	26.02	11.78	14.24	0.30
I SC CCI 7	25.89	12.12	13.77	0.41
I SC CCI 10	26.30	12.52	13.78	0.41
I SC CCI 14	26.09	11.76	14.33	0.28
I SC CCI 28	26.41	12.28	14.14	0.32
Naïve SC	27.45	11.65	15.81	0.10
I SC CFA 1	26.06	12.19	13.87	0.39
I SC CFA 3	25.63	11.66	13.97	0.36

TABLE 11-continued

Tissue	r5433	18S	∂Ct	Expression
I SC CFA 7	25.99	11.63	14.36	0.28
I SC CFA 10	25.93	11.74	14.19	0.31
I SC CFA 14	26.84	12.30	14.54	0.24
I SC CFA 28	26.27	12.18	14.09	0.33
Naïve SC	27.08	11.57	15.51	0.12
I SC AXT 1	25.87	11.89	13.98	0.36
I SC AXT 3	26.33	11.88	14.45	0.26
I SC AXT 7	26.12	11.82	14.30	0.29
I SC AXT 14	25.94	12.43	13.52	0.50

[0268] The expression of 5433 mRNA in a panel of cardiovascular (CV) organ and vessel tissues are shown below as Tables 12-13. Tables 12 and 13 show results from Taqman studies with a cardiovascular organ panel showing expression of 5433 mRNA in fetal heart, and in the artery, followed by human vascular endothelial cells (Huvec shear), aorta, and vein.

TABLE 12

Tissue Type	Mean	β 2 Mean	∂∂ Ct	Expression
H/Fetal Heart/normal/BWH 4	29.34	22.34	7	7.8125
H/Heart/Normal/Atrium/MPI 1097	35.19	21.08	14.11	0
H/Heart/Normal/Ventricle/PIT 272	35.98	18.22	17.75	0
H/Heart/Normal/Ventricle/PIT 206	36.7	21.24	15.47	0
H/Heart/Normal/Ventricle/PIT 204	37.98	18.54	19.44	0
H/Heart/Normal/Ventricle/PIT 205	33.91	18.66	15.24	0.0258
H/Heart/Diseased/Ventricle/ELI 5	32.47	17.04	15.44	0.0226
H/Heart/Diseased/Ventricle/PIT 16	33.12	18.43	14.7	0.0377
H/Heart/Diseased/Ventricle/PIT 1	38.58	22.33	16.25	0
H/Heart/Diseased/Ventricle/PIT 14	37.53	20.43	17.11	0
H/Kidney/normal/NDR 171	40	21.48	18.52	0
H/Kidney/normal/NDR 179	34.98	20.59	14.39	0.0466
H/Kidney/normal/PIT 289	39.51	24.04	15.47	0
H/Kidney/normal/PIT 351	35.9	19.61	16.3	0
H/Kidney/normal/PIT 353	37.23	19.62	17.61	0
H/Kidney/HT/NDR 233	37.13	18.77	18.36	0
H/Kidney/HT/NDR 224	38.86	19.81	19.05	0
H/Kidney/HT/CHT 1176	39.17	23.31	15.87	0
H/Kidney/HT/NDR 252	38.6	23.38	15.23	0
H/Kidney/HT/CHT 762	38.72	22.11	16.61	0
H/Skeletal Muscle/Normal/PIT 915	37.59	22.68	14.9	0
H/Skeletal Muscle/Normal/PIT 685	40	23.98	16.02	0
H/Skeletal Muscle/Normal/PIT 428	40	24.3	15.7	0
H/Liver/Normal/MPI 146	33.08	18.02	15.05	0.0295
H/Liver/Normal/CHT 339	40	22.25	17.75	0
H/Liver/Normal/CHT 1237	38.8	22.14	16.66	0

TABLE 13

Tissue Type	Mean	β 2 Mean	∂∂ Ct	Expression
Coronary SMC	39.65	21.25	18.4	0
Huvec NS	35.55	20.52	15.03	0
Huvec Shear/static pooled	32.94	21.2	11.73	0.2934
H/Adipose/PIT 695	33.87	19.64	14.23	0.052
H/Internal Mam Artery/Normal/AMC 263	35.85	21.23	14.62	0
H/Internal Mam Artery/Normal/AMC 347	37.19	21.99	15.2	0
H/Internal Mam Artery/Normal/AMC350	36.63	22.11	14.53	0
H/Internal Mam Artery/Normal/AMC 352	35.27	22.08	13.2	0
H/Artery/Normal/PIT 1180	33.95	19.48	14.47	0.0442

TABLE 13-continued

Tissue Type	Mean	β 2 Mean	∂∂ Ct	Expression
H/Artery/normal/AMC 150	34.58	23.97	10.61	0.6421
H/Artery/normal/PIT 912	35.84	21.54	14.31	0
H/Artery/normal/NDR 352	33.17	20.82	12.35	0.1922
H/Aorta/Diseased/PIT 710	36.95	21.34	15.62	0
H/Aorta/Diseased/PIT 712	33.8	20.91	12.89	0.1322
H/Aorta/Diseased/PIT 732	36.96	21.7	15.27	0
H/Artery/Diseased/iliac/NDR 753	34.99	21.02	13.97	0.0623
H/Artery/Diseased/Tibial/PIT 679	34.01	20.22	13.79	0.0706
H/Vein/Normal/PIT 1010	35.17	21	14.16	0
H/Vein/Normal/NDR 239	37.19	21.04	16.15	0
H/Vein/Normal/AMC 130	36.7	20.97	15.73	0
H/Vein/Normal/AMC 131	39.55	21.52	18.02	0
H/Vein/Normal/AMC 137	35.9	20.51	15.39	0
H/Vein/Normal/AMC 153	36.05	20.52	15.53	0
H/Vein/Normal/AMC 176	35.41	20.31	15.1	0
H/Vein/Normal/AMC 177	36.9	21.48	15.42	0
H/Vein/Normal/AMC 178	36.22	21.16	15.05	0
H/Vein/Normal/AMC 182	35.52	19.84	15.68	0
H/Vein/Normal/AMC 190	34.98	21.12	13.86	0.0675
H/Vein/Normal/AMC 192	36.6	22.39	14.21	0
H/Vein/Normal/AMC 195	36.33	21.4	14.93	0
H/Vein/Normal/AMC 211	36.53	21.11	15.43	0

Human 38554, 57301 and 58324

[0269] There are over 30 families of secondary transporters, also known as solute carriers or SLC (reviewed by Berger, et al. (2000) in *The Kidney: Physiology and Pathophysiology*, eds. Seldin D W and Giebisch G., Lippincott, Williams & Wilkins, Philadelphia 1:107-138. The SLC families are classified according to the pair of molecules they move, for instance, the SLC21 and 22 families transport organic ions. The 38554, 57301 and 58324 molecules of the invention are members of the SLC21 and SLC22 families.

Human 38554

[0270] Human 38554 is represented by two sequences containing amino acid substitutions at several residues. The human 38554 sequences (SEQ ID NO:32 or SEQ ID NO:53), which are approximately 3220 and 3227 nucleotides long, respectively, including untranslated regions, contain a predicted methionine-initiated coding sequence of about 2136 nucleotides, not including the termination codon (nucleotides 338-2473 of SEQ ID NO:32; 1-2136 of SEQ ID NO:34 or 345-2480 of SEQ ID NO:53; 1-2136 of SEQ ID NO:55). The coding sequences encode 712 amino acid proteins (SEQ ID NO:33 or SEQ ID NO:54).

[0271] Human 38554 contains the following regions or other structural features (for general information regarding PFAM identifiers, PS prefix and PF prefix domain identification numbers, refer to Sonnhammer et al. (1997) *Protein* 28:405-420: twelve transmembrane domains, thirteen non-transmembrane regions, a kazal domain (PFAM Accession Number PF00050, SEQ ID NO:41) located at about amino acid residues 476 to 523 of SEQ ID NO:33 or SEQ ID NO:54; and one peroxisomal targeting signal (PSORT PTS2, SEQ ID NO:50) at about amino acids 154 to 162 of SEQ ID NO:33 (not in SEQ ID NO:54). The transmembrane domains (predicted by MEMSAT, Jones et al., (1994) *Biochemistry* 33:3038-3049) are located at about amino acids 42 to 58, 80 to 102, 111 to 128, 190 to 212, 221 to 245, 274 to 295, 354 to 373, 393 to 414, 427 to 446, 553 to 577, 588 to 612, and 641

to 664 of SEQ ID NO:33 or SEQ ID NO:54; and the non-transmembrane regions are located at about amino acids 1 to 41, 59 to 79, 103 to 110, 129 to 189, 213 to 220, 246 to 273, 296 to 353, 374 to 392, 415 to 426, 447 to 552, 578 to 587, 613 to 640, and 665 to 712 of SEQ ID NO:33 or SEQ ID NO:54.

[0272] Human 38554 also contains the following regions or other structural features: one tyrosine kinase phosphorylation site (Prosite PS00007) at about amino acids 378 to 384 of SEQ ID NO:33 or SEQ ID NO:54; thirteen protein kinase C phosphorylation sites (Prosite PS00005) at about amino acids 4 to 6, 24 to 26, 152 to 154, 264 to 266, 312 to 314, 345 to 347, 374 to 376, 388 390, 509 to 511, 512 to 514, 629 to 631, 677 to 679 and 685 to 687 of SEQ ID NO:33 or twelve protein kinase C phosphorylation sites (Prosite PS00005) at about amino acids 4 to 6, 24 to 26, 264 to 266, 312 to 314, 345 to 347, 374 to 376, 388 390, 509 to 511, 512 to 514, 629 to 631, 677 to 679 and 685 to 687 of SEQ ID NO:54; eleven casein kinase II phosphorylation sites (Prosite PS00006) located at about amino acids 4 to 7, 31 to 34, 68 to 71, 165 to 168, 264 to 267, 304 to 307, 310 to 313, 466 to 469, 485 to 488, 677 to 680, and 694 to 697 of SEQ ID NO:33 or SEQ ID NO:54; six N-glycosylation sites (Prosite PS00001) from about amino acids 146 to 149, 309 to 312, 510 to 513, 520 to 523, 533 to 536, and 692 to 695 of SEQ ID NO:33 or SEQ ID NO:54; and twelve N-myristoylation sites (Prosite PS00008) from about amino acids 82 to 87, 226 to 231, 243 to 248, 385 to 390, 406 to 411, 446 to 451, 454 to 459, 505 to 510, 525 to 530, 537 to 537, 568 to 573, and 625 to 630 of SEQ ID NO:33 or SEQ ID NO:54.

[0273] A hydropathy plot of human 38554 was performed. Polypeptides of the invention include fragments which include: all or part of a hydrophobic sequence, e.g., the sequence from about amino acid 111 to 128, from about 274 to 295, and from about 641 to 664 of SEQ ID NO:33 or SEQ ID NO:54; all or part of a hydrophilic sequence, e.g., the sequence of from about amino acid 28 to 36, from about 134 to 142, and from about 301 to 316 of SEQ ID NO:33 or SEQ ID NO:54; a sequence which includes a Cys, or a glycosylation site.

Human 57301

[0274] The human 57301 sequence (SEQ ID NO:35), which is approximately 2866 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 1659 nucleotides, not including the termination codon (nucleotides 365-2023 of SEQ ID NO:35; 1-1659 of SEQ ID NO:37). The coding sequence encodes a 553 amino acid protein (SEQ ID NO:36).

[0275] Human 57301 contains the following regions or other structural features: twelve transmembrane domains, thirteen non-transmembrane regions, and a sugar (and other) transporter domain (PFAM Accession Number PF00083) located at about amino acid residues 106 to 530 of SEQ ID NO:36. The transmembrane domains (predicted by MEMSAT, Jones et al., (1994) *Biochemistry* 33:3038-3049) are located at about amino acids 21 to 37, 151 to 167, 174 to 196, 204 to 222, 232 to 255, 263 to 279, 352 to 369, 378 to 400, 409 to 426, 436 to 455, 466 to 486 and 495 to 515 of SEQ ID NO:36; and the non-transmembrane regions at about amino acids 1 to 20, 38 to 150, 168 to 173, 197 to 203, 223 to 231, 256 to 262, 280 to 351, 370 to 377, 401 to 408, 427 to 435, 456 to 465, 487 to 494, and 516 to 553 of SEQ ID NO:36.

[0276] Human 57301 also contains the following regions or other structural features: four protein kinase C phosphorylation sites (Prosite PS00005) at about amino acids 46 to 48, 167 to 169, 282 to 284, and 289 to 291 of SEQ ID NO:36; four casein kinase II phosphorylation sites (Prosite PS00006) located at about amino acids 35 to 38, 107 to 110, 211 to 214, and 526 to 529 of SEQ ID NO:36; two cAMP/cGMP-dependent protein kinase phosphorylation sites (Prosite PS00004) located at about amino acids 405 to 408 and 536 to 539 of SEQ ID NO:36; three N-glycosylation sites (Prosite PS00001) from about amino acids 39 to 42, 56 to 59, and 102 to 105 of SEQ ID NO:36; two amidation sites (Prosite PS00009) from about amino acids 170 to 173 and 403 to 406 of SEQ ID NO:36; and eight N-myristoylation sites (Prosite PS00008) from about amino acids 155 to 160, 187 to 192, 246 to 251, 331 to 336, 431 to 436, 443 to 448, 472 to 477 and 541 to 546 of SEQ ID NO:36.

[0277] A hydropathy plot of human 57301 was performed. Polypeptides of the invention include fragments which include: all or part of a hydrophobic sequence, e.g., the sequence from about amino acid 151 to 167, from about 263 to 279, and from about 352 to 369 of SEQ ID NO:36; all or part of a hydrophilic sequence, e.g., the sequence of from about amino acid 82 to 95, from about 325 to 332, and from about 528 to 537 of SEQ ID NO:36; a sequence which includes a Cys, or a glycosylation site.

Human 58324

[0278] The human 58324 sequence (SEQ ID NO:38), which is approximately 2480 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 2157 nucleotides, not including the termination codon (nucleotides 148-2304 of SEQ ID NO:38; 1-2157 of SEQ ID NO:40). The coding sequence encodes a 719 amino acid protein (SEQ ID NO:39).

[0279] Human 58324 contains the following regions or other structural features: twelve transmembrane domains, thirteen non-transmembrane regions, and a kazal domain (PFAM Accession Number PF00050) located at about amino acid residues 502 to 549 of SEQ ID NO:39. The transmembrane domains (predicted by MEMSAT, Jones et al., (1994) *Biochemistry* 33:3038-3049) are located at about amino acids 107 to 126, 150 to 166, 173 to 193, 231 to 254, 265 to 289, 314 to 335, 372 to 391, 420 to 444, 457 to 475, 580 to 603, 614 to 635, and 667 to 691 of SEQ ID NO:39; and the non-transmembrane regions are located at about amino acids 1 to 106, 127 to 149, 167 to 172, 194 to 230, 255 to 264, 290 to 313, 336 to 371, 392 to 419, 445 to 456, 476 to 579, 604 to 613, 636 to 666, and 692 to 719 of SEQ ID NO:39.

[0280] Human 58324 also contains the following regions or other structural features: seven protein kinase C phosphorylation sites (Prosite PS00005) at about amino acids 38 to 40, 41 to 43, 75 to 77, 342 to 344, 450 to 452, 492 to 494, and 705 to 707 of SEQ ID NO:39; five casein kinase II phosphorylation sites (Prosite PS00006) located at about amino acids 11 to 14, 129 to 132, 192 to 195, 252 to 255, and 445 to 448 of SEQ ID NO:39; one tyrosine kinase site (Prosite PS00007) located at about amino acids 144 to 151 of SEQ ID NO:39; five N-glycosylation sites (Prosite PS00001) from about amino acids 294 to 297, 300 to 303, 497 to 500, 546 to 549 and 661 to 664 of SEQ ID NO:39; one amidation site (Prosite PS00009) from about amino acids 44 to 47 of SEQ ID NO:39; and eleven N-myristoylation sites (Prosite PS00008) from about amino acids 37 to 42, 92 to 97, 100 to 105, 120 to 125, 184 to 189, 216 to 221, 264 to 269, 432 to 437, 441 to 446, 531 to 536, and 553 to 558 of SEQ ID NO:39.

[0281] A hydropathy plot of human 58324 was performed. Polypeptides of the invention include fragments which include: all or part of a hydrophobic sequence, e.g., the sequence from about amino acid 173 to 193, from about 314 to 335, and from about 667 to 691 of SEQ ID NO:39; all or part of a hydrophilic sequence, e.g., the sequence from about amino acid 40 to 50, from about 194 to 201, and from about 538 to 546 of SEQ ID NO:39; a sequence which includes a Cys, or a glycosylation site.

[0282] The 38554, 57301 and 58324 proteins contain a significant number of structural characteristics in common with members of the SLC21 or 22 transporter families. As used herein, the terms "transporter," "organic ion transporter," "organic anion transporter," "SLC21 family, or SLC22 family" include secondary active transport proteins. Secondary active transporters couple the active transport of one molecule, e.g., an ion, e.g., an organic ion (e.g., an organic anion or a cation, a prostaglandin, a steroidal compound (e.g., estrone-3-sulfate), a bile acid, a drug, a neurotransmitter, a sulfated lipophilic metabolite, a glucuronidated lipophilic metabolite, a polyamine, a carnitine, or a choline) against its concentration gradient to the energy gained by concomitant transport of a second molecule, e.g., another ion (e.g., a bicarbonate ion or a dicarboxylate ion) with its concentration gradient.

[0283] The SLC21 or SLC22 families of proteins are characterized by at least one, two, three, four, five, six, seven, eight, nine, ten, eleven, preferably twelve transmembrane domains. Typically, the hydrophobic transmembrane domains anchor the transporter within a cell or organelle membrane and through coordinated allosteric movements, affect the transport function across the membrane. The non-transmembrane loops between and beyond the transmembrane domains of the transporter determine the ion binding specificity and provide the ion binding and release activity for the transporter. Some members of these families also have a transporter domain, and/or a kazal domain.

[0284] A GAP alignment of 38554 with an SLC21 family member, organic anion transporting protein 14 (OATP-F, accession number 7839587 in GenPept, corresponding to AF260704 in Genbank, SEQ ID NO:44) results in 99.7% identity between the two sequences (as determined from a matrix made by matblas from blosum62.iij). A GAP alignment of 57301 with an SLC22 family member, organic anion transporter 4 (hOAT4, SEQ ID NO:47, accession number 7707622 in GenPept, corresponding to AB026116 in Genbank) results in 51.7% identity between the two sequences (as determined from a matrix made by matblas from blosum62.iij). A GAP alignment of 57301 with an SLC22 family member, renal-specific transporter (mouse RST, SEQ ID NO:48, accession number 2696709 in GenPept, corresponding to BAA23875 in GenBank) results in 71.4% identity between the two sequences (as determined from a matrix made by matblas from blosum62.iij). A GAP alignment of 58324 with an SLC21 family member, organic anion transporter (OATP-E, SEQ ID NO:49, accession number 6683743 in GenPept, corresponding to AB026116 in Genbank) results in 30% identity between the two sequences (as determined from a matrix made by matblas from blosum62.iij).

[0285] A 38554, 57301 or 58324 polypeptide can include at least one, two, three, four, five, six, seven, eight, nine, ten, eleven, preferably twelve "transmembrane domains" or regions homologous with a "transmembrane domain".

[0286] As used herein, the term "transmembrane domain" includes an amino acid sequence of about 10 to 40 amino acid residues in length and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, e.g., at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, e.g., leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains typically have alpha-helical structures and are described in, for example, Zagotta, W. N. et al., (1996) *Annual Rev. Neurosci.* 19:235-263, the contents of which are incorporated herein by reference.

[0287] In a preferred embodiment, a 38554, 57301 or 58324 polypeptide or protein has at least one, two, three, four, five, six, seven, eight, nine, ten, eleven, preferably twelve "transmembrane domains" or regions which include at least about 12 to 35 more preferably about 14 to 30 or 15 to 25 amino acid residues each and have at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "transmembrane domain," e.g., the transmembrane domains of human 38554, 57301 or 58324 (e.g., residues 42 to 58, 80 to 102, 111 to 128, 190 to 212, 221 to 245, 274 to 295, 354 to 373, 393 to 414, 427 to 446, 553 to 577, 588 to 612, and 641 to 664 of SEQ ID NO:33 or SEQ ID NO:54; residues 21 to 37, 151 to 167, 174 to 196, 204 to 222, 232 to 255, 263 to 279, 352 to 369, 378 to 400, 409 to 426, 436 to 455, 466 to 486 and 495 to 515 of SEQ ID NO:36; or residues 107 to 126, 150 to 166, 173 to 193, 231 to 254, 265 to 289, 314 to 335, 372 to 391, 420 to 444, 457 to 475, 580 to 603, 614 to 635, and 667 to 691 of SEQ ID NO:39). The transmembrane domains in 38554, 57301 and 58324 can be seen in hydropathy plots as regions of about 15 to 25 amino acids where the hydropathy trace is mostly above the horizontal line.

[0288] To identify the presence of a "transmembrane" domain in a 38554, 57301 or 58324 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be analyzed by a transmembrane prediction method that predicts the secondary structure and topology of integral membrane proteins based on the recognition of topological models (MEMSAT, Jones et al., (1994) *Biochemistry* 33:3038-3049).

[0289] A 38554, 57301 or 58324 polypeptide can include at least one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, preferably thirteen "non-transmembrane regions." As used herein, the term "non-transmembrane region" includes an amino acid sequence not identified as a transmembrane domain. The non-transmembrane regions in 38554 are located at about amino acids 1 to 41, 59 to 79, 103 to 110, 129 to 189, 213 to 220, 246 to 273, 296 to 353, 374 to 392, 415 to 426, 447 to 552, 578 to 587, 613 to 640, and 665 to 712 of SEQ ID NO:33 or SEQ ID NO:54. The non-transmembrane regions in 57301 are located at about amino acids 1 to 20, 38 to 150, 168 to 173, 197 to 203, 223 to 231, 256 to 262, 280 to 351, 370 to 377, 401 to 408, 427 to 435, 456 to 465, 487 to 494, and 516 to 553 of SEQ ID NO:36. The non-transmembrane regions in 58324 are located at about amino acids 1 to 106, 127 to 149, 167 to 172, 194 to 230, 255 to 264, 290 to 313, 336 to 371, 392 to 419, 445 to 456, 476 to 579, 604 to 613, 636 to 666, and 692 to 719 of SEQ ID NO:39.

[0290] The non-transmembrane regions of 38554, 57301 or 58324 include at least one, two, three, four, five, six, preferably seven cytoplasmic regions. When located at the N-terminus, the cytoplasmic region is referred to herein as the "N-terminal cytoplasmic domain." As used herein, an "N-ter-

minal cytoplasmic domain" includes an amino acid sequence having about 1 to 300, preferably about 1 to 250, preferably about 1 to 200, more preferably about 1 to 150, or even more preferably about 1 to 110 amino acid residues in length and is located inside of a cell or within the cytoplasm of a cell. The C-terminal amino acid residue of an "N-terminal cytoplasmic domain" is adjacent to an N-terminal amino acid residue of a transmembrane domain in a 38554, 57301 or 58324 protein. For example, an N-terminal cytoplasmic domain is located at about amino acid residues 1 to 41 of SEQ ID NO:33 or SEQ ID NO:54, 1 to 20 of SEQ ID NO:36, and 1 to 106 of SEQ ID NO:39.

[0291] In a preferred embodiment, a 38554, 57301 or 58324 polypeptide or protein has an N-terminal cytoplasmic domain or a region which includes at least about 5, preferably about 10 to 200, and more preferably about 15 to 110 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with an "N-terminal cytoplasmic domain," e.g., the N-terminal cytoplasmic domain of human 38554, 57301 or 58324 (e.g., residues 1 to 41 of SEQ ID NO:33 or SEQ ID NO:54, 1 to 20 of SEQ ID NO:36, and 1 to 106 of SEQ ID NO:39).

[0292] In another embodiment, a cytoplasmic region of a 38554, 57301 or 58324 protein can include the C-terminus and can be a "C-terminal cytoplasmic domain," also referred to herein as a "C-terminal cytoplasmic tail." As used herein, a "C-terminal cytoplasmic domain" includes an amino acid sequence having a length of at least about 15, preferably about 20 to 60, more preferably about 25 to 55 amino acid residues and is located inside of a cell or within the cytoplasm of a cell. The N-terminal amino acid residue of a "C-terminal cytoplasmic domain" is adjacent to a C-terminal amino acid residue of a transmembrane domain in a 38554, 57301 or 58324 protein. For example, a C-terminal cytoplasmic domain is located at about amino acid residues 665 to 712 of SEQ ID NO:33 or SEQ ID NO:54, 516 to 553 of SEQ ID NO:36, and 692 to 719 of SEQ ID NO:39.

[0293] In a preferred embodiment, a 38554, 57301 or 58324 polypeptide or protein has a C-terminal cytoplasmic domain or a region which includes at least about 5, preferably about 15 to 60, and more preferably about 25 to 55 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a C-terminal cytoplasmic domain," e.g., the C-terminal cytoplasmic domain of human 38554, 57301 or 58324 (e.g., residues 665 to 712 of SEQ ID NO:33 or SEQ ID NO:54, 516 to 553 of SEQ ID NO:36, and 692 to 719 of SEQ ID NO:39).

[0294] In another embodiment, a 38554, 57301 or 58324 protein includes at least one, two, three, four, preferably five cytoplasmic loops. As used herein, the term "loop" includes an amino acid sequence that resides outside of a phospholipid membrane, having a length of at least about 4, preferably about 5 to 150, more preferably about 6 to 120 amino acid residues, and has an amino acid sequence that connects two transmembrane domains within a protein or polypeptide. Accordingly, the N-terminal amino acid of a loop is adjacent to a C-terminal amino acid of a transmembrane domain in a 38554, 57301 or 58324 molecule, and the C-terminal amino acid of a loop is adjacent to an N-terminal amino acid of a transmembrane domain in a 38554, 57301, or 58324 molecule. As used herein, a "cytoplasmic loop" includes a loop located inside of a cell or within the cytoplasm of a cell. For example, a "cytoplasmic loop" can be found at about amino acid residues 103 to 110, 213 to 220, 296 to 353, 415 to 426, and 578 to 587 of SEQ ID NO:33 or SEQ ID NO:54; 168 to 173, 223 to 231, 280 to 351, 401 to 408, and 456 to 465 of SEQ ID NO:36; 167 to 172, 255 to 264, 336 to 371, 445 to 456, and 604 to 613 of SEQ ID NO:39.

[0295] In a preferred embodiment, a 38554, 57301 or 58324 polypeptide or protein has a cytoplasmic loop or a region which includes at least about 4, preferably about 5 to 100, and more preferably about 6 to 80 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a cytoplasmic loop," e.g., a cytoplasmic loop of human 38554, 57301 or 58324 (e.g., residues 103 to 110, 213 to 220, 296 to 353, 415 to 426, and 578 to 587 of SEQ ID NO:33 or SEQ ID NO:54; 168 to 173, 223 to 231, 280 to 351, 401 to 408, and 456 to 465 of SEQ ID NO:36; 167 to 172, 255 to 264, 336 to 371, 445 to 456, and 604 to 613 of SEQ ID NO:39).

[0296] In another embodiment, a 38554, 57301 or 58324 protein includes at least one, two, three, four, five, preferably six non-cytoplasmic loops. As used herein, a "non-cytoplasmic loop" includes an amino acid sequence located outside of a cell or within an intracellular organelle. Non-cytoplasmic loops include extracellular domains (i.e., outside of the cell) and intracellular domains (i.e., within the cell). When referring to membrane-bound proteins found in intracellular organelles (e.g., mitochondria, endoplasmic reticulum, peroxisomes microsomes, vesicles, endosomes, and lysosomes), non-cytoplasmic loops include those domains of the protein that reside in the lumen of the organelle or the matrix or the intermembrane space. For example, a "non-cytoplasmic loop" can be found at about amino acid residues 59 to 79, 129 to 189, 246 to 273, 374 to 392, 447 to 452, and 613 to 640 of SEQ ID NO:33 or SEQ ID NO:54; 38 to 150, 197 to 203, 256 to 262, 370 to 377, 427 to 435, and 487 to 494 of SEQ ID NO:36; 127 to 149, 194 to 230, 290 to 313, 392 to 419, 476 to 579, and 636 to 666 of SEQ ID NO:39.

[0297] In a preferred embodiment, a 38554, 57301 or 58324 polypeptide or protein has at least one non-cytoplasmic loop or a region which includes at least about 4, preferably about 5 to 150, more preferably about 6 to 120 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "non-cytoplasmic loop," e.g., at least one non-cytoplasmic loop of human 38554, 57301, or 58324 (e.g., residues 59 to 79, 129 to 189, 246 to 273, 374 to 392, 447 to 452, and 613 to 640 of SEQ ID NO:33 or SEQ ID NO:54; 38 to 150, 197 to 203, 256 to 262, 370 to 377, 427 to 435, and 487 to 494 of SEQ ID NO:36; 127 to 149, 194 to 230, 290 to 313, 392 to 419, 476 to 579, and 636 to 666 of SEQ ID NO:39).

[0298] A 38554, 57301 or 58324 protein or polypeptide can include a "transporter domain" or a region homologous to a "transporter domain." As used herein, the term "transporter domain" includes an amino acid sequence of about 20 to 250 amino acid residues in length, resides in a non-cytoplasmic loop and participates in the transport of a molecule; e.g. an ion, (e.g., an organic anion or cation, a hormone or a metabolite) across a membrane, e.g. a cell or organelle membrane and can have a bit score (PSI-BLAST) for the alignment of the sequence to a transporter domain of at least 80. Preferably, a transporter domain includes at least about 30 to 225 amino acids, more preferably about 35 to 215 amino acid residues, or about 40 to 195 amino acids and has a bit score for the alignment of the sequence to a transporter domain (PSI-BLAST) of at least 100, 120, 135 or greater. The transporter domain of 38554 and 58324 is homologous to ProDom fam-

ily PD005488 ("Transporter Protein Transmembrane Transport Similar Matrin F/G Organic Anion Sodium-Independent;" SEQ ID NO:43, ProDomain Release 2000.1; see also ProDomain No. PD005488, Release 1999.2). An alignment of this domain of 38554 (amino acids 476 to 667 of SEQ ID NO:33 or SEQ ID NO:54) with PD005488 resulted in 44% identity as determined by PSI-BLAST. An alignment of this domain of 58324 (amino acids 502 to 672 of SEQ ID NO:39) with PD005488 resulted in 32% identity as determined by PSI-BLAST. The transporter domain of 57301 is homologous to ProDom family PD151320 ("Organic Transporter-like Transport Protein Renal Anion Transporter Cationic Kidney-Specific Solute," SEQ ID NO:46, ProDomain Release 1999.2). An alignment of this region (amino acids 102 to 145 of SEQ ID NO:36) with PD151320 resulted in 56% identity as determined by PSI-BLAST.

[0299] In a preferred embodiment a 38554, 57301 or 58324 polypeptide or protein has a "transporter domain" or a region which includes at least about 30 to 225 amino acids, more preferably about 35 to 215 amino acid residues, or about 40 to 195 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "transporter domain," e.g., the transporter domain of human 38554, 57301 or 58324, (e.g., residues 476 to 667 of SEQ ID NO:33 or SEQ ID NO:54, 102 to 145 of SEQ ID NO:36, or 502 to 672 of SEQ ID NO:39).

[0300] For further identification of a transporter domain in a 38554, 57301, or 58324 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against a database of domains, e.g., the ProDom database (Corpet et al. (1999), *Nucl. Acids Res.* 27:263-267). The ProDom protein domain database consists of an automatic compilation of homologous domains. Current versions of ProDom are built using recursive PSI-BLAST searches (Altschul S F et al. (1997) *Nucleic Acids Res.* 25:3389-3402; Gouzy et al. (1999) *Computers and Chemistry* 23:333-340) of the SWISS-PROT 38 and TREMBL protein databases. The database automatically generates a consensus sequence for each domain. A BLAST search was performed against the HMM database resulting in the identification of a "Organic Transporter-like Transport Protein Renal Anion Transporter Cationic Kidney-Specific Solute" domain in the amino acid sequence of human 57301 at about residues 102 to 145 of SEQ ID NO:36 and a "Transporter Protein Transmembrane Transport Similar Matrin F/G Organic Anion Sodium-Independent" domain in the amino acid sequence of human 38554 and 58324 at about residues 476 to 667 of SEQ ID NO:33 or SEQ ID NO:54, or 502 to 672 of SEQ ID NO:39.

[0301] A 57301 polypeptide can further include a "sugar (and other) transporter domain" or regions homologous with a "sugar (and other) transporter domain" (SEQ ID NO:45, PFAM Accession Number PF00083). As used herein, the term "sugar (and other) transporter domain" includes an amino acid sequence of about 420 to 440 amino acid residues in length and transports molecules, e.g., ions, sugars or metabolites. An alignment of the sugar (and other) transporter domain (amino acids 106 to 530 of SEQ ID NO:36) of human 57301 with a consensus amino acid sequence (SEQ ID NO:45) derived from a hidden Markov model yields a bit score of −15.3.

[0302] Sugar (and other) transporter domains can have sequences similar to three Prosite signature sequences (two copies of PS00216 and one copy of PS00217). A sequence similar to copy one of the first Prosite signature sequence (PS00216, [LIVMSTAG]-[LIVMFSAG]-x(2)-[LIVMSA]-[DE]-x-[LIVMFYWA]-G-R-[RK]-x(4,6)-[GSTA], SEQ ID NO:51), with a mismatch at only the first residue of the consensus, is located about between the second and third transmembrane domains of the human 57301 polypeptide and can be found at about amino acids 163 to 179 of SEQ ID NO:36. A sequence similar to copy two of the first Prosite signature sequence, with a mismatch of only an S instead of the [DE], is located about between the eighth and ninth transmembrane domains of the human 57301 polypeptide and can be found at about amino acids 396 to 411 of SEQ ID NO:36. These signature sequences are involved in the conformational change required for transport. A sequence similar to the second Prosite signature sequence (PS00217, [LIVMF]-x-G-[LIVMFA]-x(2)-G-x(8)-[LIFY]-x(2)-[EQ]-x(6)-[RK], SEQ ID NO:52), with a conserved substitution of an A for the first G, a one amino acid insertion after the fourth residue of the consensus, and only one amino acid between the [LIFY] and the [EQ], is located about the end of the fourth and in the loop before the fifth transmembrane domain of the human 57301 polypeptide and can be found at about amino acids 205 to 230 of SEQ ID NO:36. In the above conserved motifs, and other motifs described herein, the standard IUPAC one-letter code for the amino acids is used. Each element in the pattern is separated by a dash (-); square brackets ([ ]) indicate the particular residues that are accepted at that position; x indicates that any residue is accepted at that position; and numbers in parentheses (( )) indicate the number of residues represented by the accompanying amino acid.

[0303] A 38554 or 58324 molecule can include a kazal domain or regions homologous with a "kazal domain" (PFAM Accession Number PF00050, SEQ ID NO:41. As used herein, the term "kazal domain" includes an amino acid sequence of about 45 to 55 amino acid residues in length and is characterized by the pattern of cysteine residues, required for disulfide bonding into a specific structure used for contact with the substrate. The kazal domain (HMM) has been assigned the SMART identifier kazal (SEQ ID NO:42). An alignment of the kazal domain (amino acids 476 to 523 of SEQ ID NO:33 or SEQ ID NO:54, or 502 to 549 of SEQ ID NO:39) of human 38554 or 58324, respectively, with a consensus amino acid sequence (SEQ ID NO:41) derived from a hidden Markov model yields a bit score of −7.7 and −13.8, respectively. An alignment of the kazal domain (amino acids 475 to 523 of SEQ ID NO:33 or SEQ ID NO:54) of human 38554 with a SMART consensus amino acid sequence (SEQ ID NO:42) derived from modular architecture analysis yields a bit score of −1.8.

[0304] To identify the presence of a "sugar (and other) transporter" domain or a "kazal" domain in a 38554, 57301, or 58324 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against the Pfam database of HMMs (e.g., the Pfam database, release 2.1) using the default parameters. For example, the hmmsf program, which is available as part of the HMMER package of search programs, is a family specific default program for MILPAT0063 and a score of 15 is the default threshold score for determining a hit. Alternatively, the threshold score for determining a hit can be lowered (e.g., to 8 bits). A description of the Pfam database can be found in Sonhammer et al. (1997) *Proteins* 28:405-420 and a detailed description of HMMs can be found, for example, in Gribskov et al. (1990) *Meth. Enzy-*

mol. 183:146-159; Gribskov et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:4355-4358; Krogh et al. (1994) *J. Mol. Biol.* 235:1501-1531; and Stultz et al. (1993) *Protein Sci.* 2:305-314, the contents of which are incorporated herein by reference. A search was performed against the HMM database resulting in the identification of a "sugar (and other) transporter" domain in the amino acid sequence of human 57301 at about residues 106 to 530 of SEQ ID NO:36; a "kazal" domain in the amino acid sequence of human 38554 at about residues 476 to 523 of SEQ ID NO:33 or SEQ ID NO:54; and a "kazal" domain in the amino acid sequence of human 58324 at about residues 502 to 549 of SEQ ID NO:39.

[0305] An additional method to identify the presence of a "kazal" domain in a 38554 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against a SMART database (Simple Modular Architecture Research Tool) of HMMs as described in Schultz et al. (1998), *Proc. Natl. Acad. Sci. USA* 95:5857 and Schultz et al. (2000) *Nucl. Acids Res* 28:231. The database contains domains identified by profiling with the hidden Markov models of the HMMer2 search program (R. Durbin et al. (1998) *Biological sequence analysis: probabilistic models of proteins and nucleic acids*. Cambridge University Press). The database also is extensively annotated and monitored by experts to enhance accuracy. A search was performed against the HMM database resulting in the identification of a "kazal" domain in the amino acid sequence of 38554 at about residues 475 to 523 of SEQ ID NO:33 or SEQ ID NO:54.

[0306] A 38554, 57301 or 58324 family member can include at least one, two, three, four, five, six, seven, eight, nine, ten, eleven, preferably twelve transmembrane domains; at least one, two, three, four, five, six, preferably seven cytoplasmic regions, including N- and C-terminal cytoplasmic domains and at least one, two, three, four, preferably five cytoplasmic loops; and at least one, two, three, four, five, preferably six non-cytoplasmic loops. Additionally, a 38554 or a 58324 family member can include at least one kazal domain and a 57301 family member can include a sugar (and other) transporter domain. A 38554 family member can further include at least one at least one peroxisomal targeting signal (PSORT PTS2).

[0307] Furthermore, a 38554 family member can include at least one, five, ten, preferably thirteen protein kinase C phosphorylation sites (PS00005); at least one, three, six, and preferably eleven casein kinase II phosphorylation sites (PS00006); at least one, two, four, and preferably six N-glycosylation sites (PS00001); at least one tyrosine kinase phosphorylation site (PS00007); and at least one, five, ten, and preferably twelve N-myristoylation sites (PS00008). A 57301 family member can include at least one, two, three, preferably four protein kinase C phosphorylation sites (PS00005); at least one, two, three, and preferably four casein kinase II phosphorylation sites, (PS00006); at least one, two, preferably three N-glycosylation sites (PS00001); at least one, preferably two cAMP- and cGMP-dependent protein kinase phosphorylation sites (PS00004); at least one, preferably two amidation sites (PS00009); and at least one, two, four, and preferably eight N-myristoylation sites (PS00008). Furthermore, a 58324 family member can include at least one, two, four, preferably seven protein kinase C phosphorylation sites (PS00005); at least one, two, three, four, and preferably five casein kinase II phosphorylation sites (PS00006); at least one, two, four, and preferably five N-glycosylation sites (PS00001); at least one tyrosine kinase phosphorylation site (PS00007); at least one amidation site (PS00009); and at least one, three, seven and preferably eleven N-myristoylation sites (PS00008).

[0308] As the 38554, 57301 or 58324 polypeptides of the invention can modulate 38554-, 57301- or 58324-mediated activities, they can be useful for developing novel diagnostic and therapeutic agents for transporter-associated or other 38554-, 57301- or 58324-associated disorders, as described below.

[0309] The SLC21 and SLC22 families are polyspecific transporters of organic ions. Some members of the SLC21 family transport organic anions, others transport prostaglandins. Some members of the SLC22 family transport organic anions, others transport organic cations, and some may transport either type of ion. They participate in activities as diverse as intestinal or hepatic absorption of metabolites, renal reabsorption of cations or excretion of cations. Members of these families also transport a wide variety of drugs and xenobiotics, many of which are harmful to the body. In addition, organic ion transporters are responsible for the transport of the metabolites of most lipophilic compounds, e.g., sulfate and glucuronide conjugates (Moller, J. V. and Sheikh, M. I. (1982) *Pharmacol Rev.* 34:315-358; Pritchard, J. B. and Miller, D. S. (1993) *Physiol. Rev.* 73:765-796; Ullrich, K. J. (1997) *J. Membr. Biol.* 158:95-107; Ullrich, K. J. and Rumrich, G. (1993) *Clin. Investig.* 71:843-848; Petzinger, E. (1994) *Rev. Physiol. Biochem. Pharmacol.* 123:47-211).

[0310] Proper function of members of these families is important for many physiological processes. At the cellular level, aberrant or deficient organic ion transporter activity can detrimentally affect functions such as cellular proliferation, growth, differentiation, or migration. At the tissue level, aberrant or deficient organic ion transporter activity can detrimentally affect inter- or intra-cellular communication; or musculoskeletal function. At the organ level, aberrant or deficient organic ion transporter activity can detrimentally affect kidney, liver or cardiac function. At the organism level, aberrant or deficient organic ion transporter activity can detrimentally affect systemic responses, such as nervous system responses, hormonal responses (e.g., insulin response), or immune responses; and protection of cells from toxic compounds (e.g., carcinogens, toxins, or mutagens).

[0311] As used herein, a "38554, 57301 or 58324 activity", "biological activity of 38554, 57301 or 58324" or "functional activity of 38554, 57301 or 58324", refers to an activity exerted by a 38554, 57301 or 58324 protein, polypeptide or nucleic acid molecule on e.g., a 38554-, 57301- or 58324-responsive cell or on a 38554, 57301 or 58324 substrate, e.g., a protein substrate, as determined in vivo or in vitro. In one embodiment, a 38554, 57301 or 58324 activity is a direct activity, such as an association with a 38554, 57301 or 58324 target molecule. A "target molecule" or "binding partner" is a molecule with which a 38554, 57301 or 58324 protein binds or interacts in nature. In an exemplary embodiment, 38554, 57301 or 58324 is a transporter, e.g., an SLC21 or 22 family organic ion transporter, and thus binds to or interacts in nature with a molecule, e.g., an organic ion.

[0312] A 38554, 57301 or 58324 activity can also be an indirect activity, e.g., a cellular signaling activity mediated by interaction of the 38554, 57301 or 58324 protein with a 38554, 57301 or 58324 receptor. Based on the above-described sequence structures and similarities to molecules of known function, the 38554, 57301 or 58324 molecules of the

present invention have similar biological activities as SLC21 or 22 family members. For example, the 38554, 57301 or 58324 proteins of the present invention can have one or more of the following activities: (1) the ability to reside within a membrane; (2) the ability to interact with a substrate or target molecule; (3) the ability to transport a substrate or target molecule, e.g., an ion, e.g., an organic ion (e.g., an organic anion, an organic cation, a prostaglandin, a steroidal compound (e.g., estrone-3-sulfate), a bile acid, a drug, a neurotransmitter, a sulfated lipophilic metabolite, a glucuronidated lipophilic metabolite, a polyamine, a carnitine, or a choline) across a membrane; (4) the ability to transport a second substrate or target molecule, e.g., an ion, (e.g., a bicarbonate ion or a dicarboxylate ion), across a membrane; (5) the ability to interact with and/or modulate a second non-transporter protein; (6) the ability to modulate cellular signaling and/or gene transcription (e.g., either directly or indirectly); (7) the ability to protect cells and/or tissues from organic ions; (8) the ability to protect cells and/or tissues from organic anions; (9) the ability to modulate hormonal responses; (10) the ability to modulate metabolism; and (11) the ability to modulate excretion.

[0313] The 38554, 57301 or 58324 molecules of the invention can modulate the activities of cells in tissues where they are expressed. For example in TaqMan analysis, 38554 mRNA is expressed at high levels in human brain cortex and hypothalamus tissue and at medium levels in dorsal root ganglion, spinal cord, choroid plexus, and testes. In the neurological tissues, the expression is found on glial cells, with an epithelial cell similarity in choroid plexus. Expression of 38554 mRNA in monkey and rodent neurological tissues confirms the expression found in human neurological tissues. Regulation of expression is found in rodent dorsal root ganglion after axotomy. Also for example, 57301 mRNA is expressed at high levels in kidney and 58324 mRNA is expressed at small levels in hemangioma tissue. Accordingly, the 38554, 57301 or 58324 molecules of the invention can act as therapeutic or diagnostic agents for one or more of a pain disorder, a nervous system disorder, an immune, e.g., inflammatory disorder, a testicular disorder, a kidney disorder, or an angiogenesis disorder, as well as disorders in tissues where 38554 molecules are expressed at lower levels as described below. Small amounts of 38554 expression were found in normal artery, human umbilical vein endothelial cells, hemangioma tissue, tissue from heart undergoing congestive heart failure, and kidney. Trace amounts of 32468 expression were found in salivary glands, normal colon, colon tumor, normal lung, normal tonsil, mammary gland and pancreas.

[0314] The 38554, 57301 or 58324 molecules of the invention can play an important role in pain disorders. In addition, the 38554 molecules of the invention can be used to treat and/or diagnose pain disorders in part because 38554 mRNA is expressed in glial cells, cells with important roles in neuropathic pain and/or because 38554 expression is regulated in the dorsal root ganglion after axotomy.

[0315] The 38554 molecules can be used to treat neurological disorders in part because the 38554 mRNA is expressed in the brain cortex, hypothalamus tissue, dorsal root ganglion, spinal cord, and choroid plexus.

[0316] The 38554 molecules of the invention can be used to treat and/or diagnose a variety of immune, e.g., inflammatory disorders in part because the 38554 mRNA is expressed in the choroid plexus. The choroid plexus is responsible for the secretion of the cerebral fluid and is involved in inflammatory responses

[0317] The 38554 molecules can be used to treat testicular disorders in part because the 38554 mRNA is expressed in the testis. The blood-testis barrier is analogous to the blood-brain barrier in the physiology of seminiferous tubules and maturation of spermatozoa as they develop into spermatids. Transporter molecules, such as 38554, can play a role in the maintenance of this barrier and supply of ions to the developing spermatids.

[0318] The 57301 molecules can be used to treat renal disorders in part because the 57301 mRNA is expressed in the kidney.

[0319] The 38554 and 58324 molecules can be used to treat angiogenic disorders in part because the 38554 mRNA is expressed in normal artery, human umbilical vein endothelial cells, hemangioma tissue, and tissue from heart undergoing congestive heart failure, and 58324 mRNA is expressed in hemangioma tissue.

[0320] Thus, the 38554, 57301 or 58324 molecules can act as novel diagnostic targets and therapeutic agents for controlling one or more of a pain disorder, a nervous system disorder, an immune, e.g., inflammatory disorder, a testicular disorder, a kidney disorder, an angiogenesis disorder, as described above, or other organic ion transport, organic ion absorption or excretion, inter- or intra-cellular signaling, and/or hormonal response disorders.

Gene Expression Analysis

[0321] Human 38554, 57301 and 58324 expression was measured by TaqMan® quantitative PCR (Perkin Elmer Applied Biosystems) in cDNA prepared from a variety of normal and diseased (e.g., cancerous) human tissues or cell lines.

[0322] This analysis found 38554 mRNA expression at high levels in human brain cortex and hypothalamus tissue, at medium levels in dorsal root ganglion, spinal cord, choroid plexus, and testes, at small levels in normal artery, human umbilical vein endothelial cells, hemangioma tissue, tissue from heart undergoing congestive heart failure, and kidney; and at trace levels in salivary glands, normal colon, colon tumor, normal lung, normal tonsil, mammary gland and pancreas; 57301 mRNA expression at high levels in kidney; and 58324 mRNA expression at small levels in hemangioma tissue.

Expression of 38554 in Rodent Tissue and Rodent Pain Models

[0323] To study the expression of 38554 in animal models of pain, rats were subjected to the following procedures: ligation of the sciatic nerve to produce chronic constriction injury (Bennett G J & Xie Y K, 1988; Pain 33; 87-107), plantar injection of complete Freund's adjuvant (Stein C, Millan M J and Herz A, 1988; Pharmacol Biochem Behav 31; 445-451) to produce inflammatory pain, or axotomy of the sciatic nerve (Curtis et al., 1994; Neuron 12; 191-204) to produce chronic pain. TaqMan® quantitative PCR (PE Applied Biosystems) to measure the expression of the rat ortholog of human 38554 in cDNA prepared from a variety of normal and diseased (e.g., pain models) tissues was performed by the same methods as for the human tissue, as described above, except 18S RNA was used as an internal

amplicon reference and reference probe; the integrity of the RNA samples following DNase I treatment was confirmed by 1.2% agarose gel electrophoresis; probes were designed by PrimerExpress software (PE Biosystems) based on the sequence of the rat 38554 gene; 200 nM of forward and reverse primers plus 100 nM probe for 18S and 900 nM forward and reverse primers plus 250 nM probe; and the Ct value of the rat 38554 gene is normalized by subtracting the Ct value of the 18S to obtain a $_{\Delta}$Ct value using the following formula: $_{\Delta}Ct = averageCt_{rat\ 38554} - averageCt_{18S}$.

[0324] The results indicated high levels of rat 38554 expression in brain and spinal cord, with low levels of rat 38554 expression in dorsal root ganglion. There was trace levels of rat 38554 expression in superior cervical ganglion, ovary and uterus. In the analysis of tissues using pain models, up-regulation of rat 38554 expression is found in rodent dorsal root ganglion after axotomy.

In Situ Hybridization

[0325] The human 38554 clone was used to make probes for in situ hybridization experiments. Expression of 38554 mRNA in monkey and rodent neurological tissues confirms the expression found by TaqMan® quantitative PCR in human neurological tissues. In the rat neurological tissues, in situ hybridization found 38554 expression in cortex, hippocampus, spinal cord, and cerebrum. In the monkey neurological tissues, in situ hybridization found 38554 expression in the cortex, choroid plexus and dorsal root ganglion. Glial cells were the specific cell type labeled in these tissues; in choroid plexus, the cells are similar to epithelial cells.

Human 55063 (NMDA-1)

[0326] The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as "human NMDA-1" or "HNMDA-1" or "55063" nucleic acid and polypeptide molecules, which are novel members of the glutamate-gated ion channel family. These novel molecules are capable of, for example, modulating a glutamate-gated ion channel mediated activity (e.g., an NMDA mediated activity) in a neural cell (e.g., in the brain and/or spinal cord). These novel molecules are capable of binding neurotransmitters, e.g., L-glutamate and glycine, and transporting ions, e.g., Ca^{2+}, across neural membranes and, thus, play a role in or function in a variety of cellular processes, e.g., mediating excitatory postsynaptic currents (e.g., long term potentiation).

[0327] The human NMDA-1 or 55063 sequence (SEQ ID NO:56), which is approximately 4197 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 3345 nucleotides, not including the termination codon (nucleotides 1-3345 of SEQ ID NO:56; 1-3345 of SEQ ID NO:58). The coding sequence encodes a 1115 amino acid protein (SEQ ID NO:57).

[0328] The HNMDA-1 amino acid sequence was aligned with the amino acid sequence of the rat NMDA-L (Accession No. 1050330; SEQ ID NO:59) amino acid sequence using the CLUSTAL W (1.74) multiple sequence alignment program to show the homology between the two proteins.

[0329] A search using the polypeptide sequence of SEQ ID NO:57 was performed against the HMM database in PFAM resulting in the identification of a potential ligand-gated ion channel family domain in the amino acid sequence of HNMDA-1 at about residues 674-952 of SEQ ID NO:57 (score=198.1).

[0330] A search using the polypeptide sequence of SEQ ID NO:57 was also performed against the HMM database in SMART resulting in the identification of a potential glutamate-gated ion channel family domain in the amino acid sequence of HNMDA-1 at about residues 565-910 of SEQ ID NO:57 (score=267.4).

[0331] The amino acid sequence of HNMDA-1 was analyzed using the program PSORT to predict the localization of the proteins within the cell. This program assesses the presence of different targeting and localization amino acid sequences within the query sequence. The results of this analysis show that HNMDA-1 may be localized to the endoplasmic reticulum, mitochondria, or nucleus.

[0332] Searches of the amino acid sequence of HNMDA-1 were further performed against the Prosite database. These searches resulted in the identification in the amino acid sequence of HNMDA-1 of a number of potential N-glycosylation sites, a potential cAMP- and cGMP-dependent protein kinase phosphorylation site, a number of potential protein kinase C phosphorylation sites, a number of potential casein kinase II phosphorylation sites, a potential tyrosine kinase phosphorylation site, a number of potential N-myristoylation sites, a number of potential amidation sites, and a potential ATP/GTP-binding site motif A (P-loop).

[0333] A MEMSAT analysis of the polypeptide sequence of SEQ ID NO:57 was also performed predicting three possible transmembrane domains in the amino acid sequence of HNMDA-1 (SEQ ID NO:57) at about residues 677-695, 748-770, and 931-951. Further analysis of the amino acid sequence of SEQ ID NO:57 (e.g., alignment with, for example, a known rat NMDA protein, SEQ ID NO:59) resulted in the identification of a fourth transmembrane domain at about amino acid residues 713-734 of SEQ ID NO:57.

[0334] As used herein, a "glutamate-gated ion channel" includes a protein or polypeptide which is a member of the ligand-gated ion channel family and is involved in binding ligands, (e.g., binding L-glutamate and glycine), and transporting ions (e.g., Ca^{2+}) across the plasma membrane of a cell (e.g., a neural cell). Glutamate-gated ion channels regulate long term potentiation in a cell and, typically, have glutamate substrate specificity. Examples of glutamate-gated ion channels include kainate, AMPA, and NMDA receptors.

[0335] As used herein, a "glutamate-gated ion channel mediated activity" includes an activity which involves a glutamate-gated ion channel in a cell (e.g., in a neural cell). Glutamate-gated ion channel mediated activities include the binding of a ligand (e.g., L-glutamine and/or glycine); the transporting of Ca^{2+} across a neural membrane; the regulation of long term potentiation; and the regulation of synapse formation underlying memory, learning, and formation of neural networks during development.

[0336] As the HNMDA-1 molecules of the present invention are glutamate-gated ion channels, they may be useful for developing novel diagnostic and therapeutic agents for glutamate-gated ion channel associated disorders. As used herein, the term "glutamate-gated ion channel associated disorder" includes a disorder, disease, or condition which is characterized by an aberrant, e.g., upregulated or downregulated, glutamate-gated ion channel mediated activity. Glutamate-gated ion channel associated disorders typically

result in, e.g., upregulated or downregulated, Ca^{2+} levels in a cell (e.g., a neural cell). Examples of glutamate-gated ion channel associated disorders include disorders associated with long term synapse potentiation, acute and chronic neurological disorders, psychiatric disorders, and neuropathic pain syndromes. Glutamate-gated ion channel associated disorders can detrimentally affect conveyance of sensory impulses from the periphery to the brain and/or conductance of motor impulses from the brain to the periphery; integration of reflexes; interpretation of sensory impulses; and emotional, intellectual (e.g., learning and memory), or motor processes.

[0337] The term "family" when referring to the polypeptide and nucleic acid molecules of the invention is intended to mean two or more polypeptides or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein. For example, the family of HNMDA-1 polypeptides comprise at least one "transmembrane domain" and preferably four transmembrane domains. As used herein, the term "transmembrane domain" includes an amino acid sequence of about 20-45 amino acid residues in length which spans the plasma membrane. More preferably, a transmembrane domain includes about at least 20, 25, 30, 35, 40, or 45 amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an alpha-helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, e.g., leucines, isoleucines, alanines, valines, phenylalanines, prolines or methionines. Transmembrane domains are described in, for example, Zagotta W. N. et al, (1996) *Annual Rev. Neurosci.* 19: 235-263, the contents of which are incorporated herein by reference. A MEMSAT analysis resulted in the identification of four transmembrane domains in the amino acid sequence of HNMDA-1 (SEQ ID NO:57) at about residues 7-28, 677-695, 748-770, and 931-951.

[0338] Accordingly, HNMDA-1 polypeptides having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with a transmembrane domain of human HNMDA-1 are within the scope of the invention.

[0339] In another embodiment of the invention features HNMDA-1 molecules which contain a signal sequence. As used herein, a "signal peptide" includes a peptide of at least about 20 amino acid residues in length which occurs at the N-terminus of secretory and integral membrane proteins and which contains at least 55% hydrophobic amino acid residues. In a preferred embodiment, a signal sequence contains at least about 15-45 amino acid residues, preferably about 20-42 amino acid residues. Signal sequences of 25-35 amino acid residues and 28-32 amino acid residues are also within the scope of the invention. As used herein, a signal sequence has at least about 40-70%, preferably about 50-65%, and more preferably about 55-60% hydrophobic amino acid residues (e.g., Alanine, Valine, Leucine, Isoleucine, Phenylalanine, Tyrosine, Tryptophan, or Proline). Such a "signal peptide", also referred to in the art as a "signal sequence", serves to direct a protein containing such a sequence to a lipid bilayer. For example, a signal sequence can be found at about amino acids 1-22 of SEQ ID NO:57 (Met1 to Ala22 of the HNMDA-1 amino acid sequence).

[0340] Accordingly, HNMDA-1 polypeptides having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with a signal sequence domain of HNMDA-1 are within the scope of the invention.

[0341] In another embodiment, an HNMDA-1 molecule of the present invention is identified based on the presence of at least one "ligand-gated ion channel family domain." As used herein, the term "ligand-gated ion channel family domain" includes a protein domain having at least about 200-400 amino acid residues, having a bit score of at least 100 when compared against a ligand-gated ion channel family domain Hidden Markov Model (HMM), and, preferably, a ligand-gated ion channel mediated activity. Preferably, a ligand-gated ion channel family domain includes a polypeptide having an amino acid sequence of about 250-400, 250-350, or more preferably, about 278 amino acid residues, a bit score of at least 160, 170, 180, 190, or more preferably about 198.1, and, preferably a ligand-gated ion channel mediated activity. To identify the presence of a ligand-gated ion channel family domain in an HNMDA-1 protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein may be searched against a database of known protein domains (e.g., the PFAM HMM database). A PFAM ligand-gated ion channel family domain has been assigned the PFAM Accession PF00060. A search was performed against the PFAM HMM database resulting in the identification of a ligand-gated ion channel family domain in the amino acid sequence of an HNMDA-1 at about residues 674-952 of SEQ ID NO:57.

[0342] Preferably, a "ligand-gated ion channel family domain" has a "ligand-gated ion channel mediated activity" as described herein. For example, a ligand-gated ion channel family domain may have the ability to bind a ligand, e.g., a neurotransmitter (e.g., acetylcholine, serotonin, glycine, glutamate, and/or GABA), on a cell (e.g., a neural cell); and the ability to regulate ion transport in a cell (e.g., Ca^{2+}, K^{+}, H^{+}, Cl^{-}, Mg^{2+} and/or Na^{+}). Accordingly, identifying the presence of a "ligand-gated ion channel family domain" can include isolating a fragment of an HNMDA-1 molecule (e.g., an HNMDA-1 polypeptide) and assaying for the ability of the fragment to exhibit one of the aforementioned ligand-gated ion channel mediated activities.

[0343] In another embodiment, an HNMDA-1 molecule of the present invention is identified based on the presence of at least one "glutamate-gated ion channel family domain." As used herein, the term "glutamate-gated ion channel family domain," also known as an "ionotropic glutamate receptor family domain," includes a protein domain having at least about 200-500 amino acid residues, having a bit score of at least 200 when compared against a glutamate-gated ion channel family domain Hidden Markov Model (HMM), and a glutamate-gated ion channel mediated activity. Preferably, a glutamate-gated ion channel family domain includes a polypeptide having an amino acid sequence of about 250-450, 300-400, 325-375, or more preferably, about 345 amino acid residues, a bit score of at least 210, 220, 230, 240, 250, 260, or more preferably about 267.4, and a glutamate-gated ion channel mediated activity. To identify the presence of a glutamate-gated ion channel family domain in an HNMDA-1 protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein may be searched against a database of known protein domains (e.g., the PFAM HMM database). A PFAM glutamate-gated

ion channel family domain has been assigned the InterPro Accession IPR001320. A search was performed against the PFAM HMM database resulting in the identification of a glutamate-gated ion channel family domain in the amino acid sequence of an HNMDA-1 at about residues 565-910 of SEQ ID NO:57.

[0344] Preferably, a "glutamate-gated ion channel family domain" has a "glutamate-gated ion channel mediated activity" as described herein. For example, a glutamate-gated ion channel family domain may have the ability to bind a ligand, e.g., L-glutamate and/or glycine, on a cell (e.g., a neural cell); and the ability to regulate Ca^{2+} transport in a cell. Accordingly, identifying the presence of a "glutamate-gated ion channel family domain" can include isolating a fragment of an HNMDA-1 molecule (e.g., an HNMDA-1 polypeptide) and assaying for the ability of the fragment to exhibit one of the aforementioned glutamate-gated ion channel mediated activities.

[0345] A description of the Pfam database can be found in Sonhammer et al. (1997) *Proteins* 28:405-420 and a detailed description of HMMs can be found, for example, in Gribskov et al. (1990) *Meth. Enzymol.* 183:146-159; Gribskov et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:4355-4358; Krogh et al. (1994) *J. Mol. Biol.* 235:1501-1531; and Stultz et al. (1993) *Protein Sci.* 2:305-314, the contents of which are incorporated herein by reference.

[0346] In a preferred embodiment, the HNMDA-1 molecules of the invention include at least one, preferably two, more preferably three, and even more preferably four transmembrane domain(s) and at least one of the following domains: a signal peptide, a ligand-gated ion channel family domain, and/or a glutamate-gated ion channel family domain.

[0347] Isolated HNMDA-1 polypeptides of the present invention have an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:57 or are encoded by a nucleotide sequence sufficiently identical to SEQ ID NO:56 or 58. As used herein, the term "sufficiently identical" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (e.g., an amino acid residue which has a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences share common structural domains or motifs and/or a common functional activity. For example, amino acid or nucleotide sequences which share common structural domains having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homology or identity across the amino acid sequences of the domains and contain at least one and preferably two structural domains or motifs, are defined herein as sufficiently identical. Furthermore, amino acid or nucleotide sequences which share at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homology or identity and share a common functional activity are defined herein as sufficiently identical.

[0348] In a preferred embodiment, an HNMDA-1 polypeptide includes at least one or more of the following domains: a transmembrane domain, a signal peptide, a ligand-gated ion channel family domain, and/or a glutamate-gated ion channel family domain, and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homologous or identical to the amino acid sequence of SEQ ID NO:57. In yet another preferred embodiment, an HNMDA-1 polypeptide includes at least one or more of the following domains: a transmembrane domain, a signal peptide, a ligand-gated ion channel family domain, and/or a glutamate-gated ion channel family domain, and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a complement of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:56 or SEQ ID NO:58. In another preferred embodiment, an HNMDA-1 polypeptide includes at least one or more of the following domains: a transmembrane domain, a signal peptide, a ligand-gated ion channel family domain, and/or a glutamate-gated ion channel family domain, and has an HNMDA-1 activity.

[0349] As used interchangeably herein, an "HNMDA-1 activity", "biological activity of HNMDA-1" or "functional activity of HNMDA-1," refers to an activity exerted by an HNMDA-1 polypeptide or nucleic acid molecule on an HNMDA-1 responsive cell or tissue, or on an HNMDA-1 polypeptide substrate, as determined in vivo, or in vitro, according to standard techniques. In one embodiment, an HNMDA-1 activity is a direct activity, such as an association with an HNMDA-1-target molecule. As used herein, a "substrate," "target molecule," or "binding partner" is a molecule with which an HNMDA-1 polypeptide binds or interacts in nature, such that HNMDA-1-mediated function is achieved. An HNMDA-1 target molecule can be a non-HNMDA-1 molecule or an HNMDA-1 polypeptide or polypeptide of the present invention. In an exemplary embodiment, an HNMDA-1 target molecule is an HNMDA-1 ligand, e.g., a glutamate-gated ion channel ligand such as L-glutamate or glycine. Alternatively, an HNMDA-1 activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the HNMDA-1 polypeptide with an HNMDA-1 ligand. The biological activities of HNMDA-1 are described herein. For example, the HNMDA-1 polypeptides of the present invention can have one or more of the following activities: (1) modulate Ca^{2+} transport across a cell membrane, (2) modulate intracellular Ca^{2+} concentration, (3) bind a ligand, e.g., L-glutamate, and/or glycine, (4) influence long term synapse potentiation, (5) modulate synapse formation, e.g., synapse formation related to memory or learning, and/or (6) modulate synapse formation related to the formation of neural networks during development.

Human 52991

[0350] Membrane transport molecules (e.g., channels/pores, permeases, and transporters) play important roles in the ability of the cell to regulate homeostasis, to grow and divide, and to communicate with other cells, e.g., to secrete and receive signaling molecules, such as hormones, reactive oxygen species, ions, neurotransmitters, and cytokines. A wide variety of human diseases and disorders are associated with defects in transporter or other membrane transport molecules, including certain types of liver disorders (e.g., due to defects in transport of long-chain fatty acids (Al Odaib et al. (1998) *New Eng. J. Med.* 339: 1752-1757)), hyperlysinemia (due to a transport defect of lysine into mitochondria (Oyanagi et al. (1986) *Inherit. Metab. Dis.* 9: 313-316), and cataracts (Wintour (1997) *Clin Exp Pharmacol Physiol* 24(1): 1-9). The present invention is based, in part, on the discovery of a novel human transporter, referred to herein as "52991"

[0351] The human 52991 sequence (SEQ ID NO:60), which is approximately 2247 nucleotides long including

untranslated regions, contains a predicted methionine-initiated coding sequence of about 1995 nucleotides (nucleotides 51-2045 of SEQ ID NO:60; 1-1995 of SEQ ID NO:62), not including the terminal codon. The coding sequence encodes a 665 amino acid protein (SEQ ID NO:61).

[0352] This mature protein form is approximately 665 amino acid residues in length (from about amino acid 1 to amino acid 665 of SEQ ID NO:61). Human 52991 contains the following regions or other structural features: a predicted Na+ dependent nucleoside transporter domain located at about amino acid residues 198-587 of SEQ ID NO:61; thirteen predicted transmembrane domains which extend from about amino acid residues 104-120, 128-144, 175-191, 201-217, 228-244, 262-282, 289-313, 336-354, 363-382, 418-442, 454-473, 528-550 and 568-586 of SEQ ID NO:61; five predicted N-glycosylation sites (PS00001) located at about amino acids 30-33, 34-37, 604-607, 610-613, and 638-641 of SEQ ID NO:61; nine predicted protein kinase C phosphorylation sites (PS00005) located at about amino acids 36-38, 100-102, 193-195, 385-387, 391-393, 523-525, 556-558, 611-613, and 643-645 of SEQ ID NO:61; four predicted casein kinase II phosphorylation sites (PS00006) located at about amino acids 50-53, 300-303, 330-333, and 589-592 of SEQ ID NO:61; one predicted tyrosine kinase phosphorylation site (PS00007) located at about amino acids 80-87 of SEQ ID NO:61; ten predicted N-myristoylation sites (PS00008) located at about amino acids 107-112, 347-352, 357-362, 413-418, 544-549, 564-569, 572-577, 584-589, 636-641, and 660-665 of SEQ ID NO:61; one predicted amidation site (PS00009) located at about amino acids 161-164 of SEQ ID NO:61; and two predicted prokaryotic membrane lipoprotein lipid attachment sites (PS00013) located at about amino acids 111-121 and 571-581 of SEQ ID NO:61.

[0353] For general information regarding PFAM identifiers, PS prefix and PF prefix domain identification numbers, refer to Sonnhammer et al. (1997) *Protein* 28:405-420.

[0354] A hydropathy plot of human 52991 was performed. Polypeptides of the invention include fragments which include: all or part of a hydrophobic sequence, e.g., the sequence from about amino acid 105 to 120, from about 420 to 440, and from about 530 to 550 of SEQ ID NO:61; all or part of a hydrophilic sequence, e.g., a sequence below the dashed line, e.g., the sequence from about amino acid 20 to 35, from about 65 to 85, and from about 380 to 390 of SEQ ID NO:61; a sequence which includes a Cys, or a glycosylation site.

[0355] The 52991 protein contains a significant number of structural characteristics in common with members of the transporter family. The term "family" when referring to the protein and nucleic acid molecules of the invention means two or more proteins or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein. As used herein, the term "transporter" includes a molecule which is involved in the movement of an ion or a biochemical molecule from one side of a lipid bilayer to the other, for example, against a preexisting concentration gradient. Transporters are usually involved in the movement of biochemical compounds which would normally not be able to cross a membrane (e.g., a protein, an ion, or other small molecule, such as ATP, signaling molecules, vitamins, and cofactors). Transporter molecules are involved in the growth, development, and differentiation of cells, in the regulation of cellular homeostasis, in the metabolism and catabolism of biochemical molecules necessary for energy production or storage, in intra- or intercellular signaling, in metabolism or catabolism of metabolically important biomolecules, and in the removal of potentially harmful compounds from the interior of the cell. Examples of transporters include GSH transporters, ATP transporters, and fatty acid transporters. As transporters, the transporter molecules of the present invention provide novel diagnostic targets and therapeutic agents to control transporter-associated disorders.

[0356] As used herein, a "52991 activity", "biological activity of 52991" or "functional activity of 52991", refers to an activity exerted by a 52991 protein, polypeptide or nucleic acid molecule on e.g., a 52991-responsive cell or on a 52991 substrate, e.g., a lipid or protein substrate, as determined in vivo or in vitro. In one embodiment, a 52991 activity is a direct activity, such as an association with a 52991 target molecule. A "target molecule" or "binding partner" is a molecule with which a 52991 protein binds or interacts in nature, e.g., a molecule to which the 52991 protein transports across a biological membrane. A 52991 activity can also be an indirect activity, e.g., a cellular signaling activity mediated by interaction of the 52991 protein with a 52991 ligand. For example, the 52991 proteins of the present invention can have one or more of the following activities: 1) modulate the import and export of molecules from cells, e.g., sugars, amino acids and possibly other metabolites, 2) modulate intra- or intercellular signaling, 3) modulate removal of potentially harmful compounds from the cell, or facilitate the compartmentalization of these molecules into a sequestered intracellular space (e.g., the peroxisome), and 4) modulate transport of molecules across membranes, e.g., the plasma membrane, or the membrane of a mitochondrion, a peroxisome, a lysosome, the endoplasmic reticulum, the nucleus, or a vacuole and 5) the ability to antagonize or inhibit, competitively or non-competitively, any of 1-4. Therefore, the 52991 protein may play a role in the transport of molecules into cells or across membranes in cells or organelles that lack such molecules or alternatively in the transport of molecules across membranes from cells or organelles that have an excess of such molecules.

[0357] The 52991 transporter protein has similarities to previously characterized sodium nucleoside cotransporters. (Huang, Q. Q., et al. (1994) *J Biol Chem* 269(27):17757-60; Ritzel, M. W. et al. (1997) *Am J Phiosiol* 272(2 Pt 1):C707-14; Pajor, A. M. *Biochim Biosphys Acta* (1998) 141(1):266-9; Che, M et al. (1995) *J Biol Chem* 270(23):13596-9; Yao, S. Y., et al. (1996) *Mol Pharmacol* 50(6):127-35). Thus, the 52991 transporter may play a role similar to that of such known cotransporters in transporting nucleosides or nucleoside analogs. More specifically, the 52991 transporter may be involved in the intestinal absorption and/or the renal handling of pyrimidine nucleosides, such a thymidine and uridine, or pyrimidine analogs, such as AZT, used to treat transporter-associated disorders. Therefore, regulation of 52991 transporter activity may be an important strategy in controlling transporter-associated disorders associated with the inhibition or stimulation of 52991 transporter activity.

[0358] A 52991 polypeptide can include a "Na+ dependent nucleoside transporter domain" or regions homologous with a "Na+ dependent nucleoside transporter domain". As used herein, the term "Na+ dependent nucleoside transporter domain" includes an amino acid sequence of about 200-500 amino acid residues in length and having a bit score for the alignment of the sequence to the transporter domain (HMM)

of at least 8. Preferably, a transporter domain includes at least about 300-450 amino acids, more preferably about 350-425 amino acid residues, or about 375-400 amino acids and has a bit score for the alignment of the sequence to the transporter domain (HMM) of at least 16, 50, 100, 200, 300, 400, 500 or greater. The transporter domain (HMM) has been assigned the PFAM Accession PF01773. The transporter domain (amino acids 198 to 587 of SEQ ID NO:61) of human 52991 aligns with a consensus amino acid sequence (SEQ ID NO:63) derived from a hidden Markov model.

[0359] In a preferred embodiment 52991 polypeptide or protein has a "Na+ dependent nucleoside transporter domain" or a region which includes at least about 200-500 more preferably about 300-450 or 375-400 amino acid residues and has at least about 60%, 70%, 80%, 90%, 95%, 99%, or 100% homology with a "Na+ dependent nucleoside transporter domain," e.g., the Na+ dependent nucleoside transporter domain of human 52991 (e.g., amino acid residues 198-587 of SEQ ID NO:61).

[0360] To identify the presence of a "transporter" domain in a 52991 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against a database of HMMs (e.g., the Pfam database, release 2.1) using the default parameters. For example, the hmmsf program, which is available as part of the HMMER package of search programs, is a family specific default program for MILPAT0063 and a score of 15 is the default threshold score for determining a hit. Alternatively, the threshold score for determining a hit can be lowered (e.g., to 8 bits). A description of the Pfam database can be found in Sonhammer et al., (1997) *Proteins* 28(3):405-420 and a detailed description of HMMs can be found, for example, in Gribskov et al., (1990) *Meth. Enzymol.* 183:146-159; Gribskov et al., (1987) *Proc. Natl. Acad. Sci. USA* 84:4355-4358; Krogh et al., (1994) *J. Mol. Biol.* 235:1501-1531; and Stultz et al., (1993) *Protein Sci.* 2:305-314, the contents of which are incorporated herein by reference.

[0361] A search was performed against the HMM database resulting in the identification of a "Na+ dependent nucleoside transporter" domain in the amino acid sequence of human 52991 at about residues 198-587 of SEQ ID NO:61.

[0362] For further identification of domains in a 52991 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against a database of domains, e.g., the ProDom database (Corpet et al. (1999), Nucl. Acids Res. 27:263-267). The ProDom protein domain database consists of an automatic compilation of homologous domains. Current versions of ProDom are built using recursive PSI-BLAST searches (Altschul S F et al. (1997) *Nucleic Acids Res.* 25:3389-3402; Gouzy et al. (1999) 23:333-340) of the SWISS-PROT 38 and TREMBL protein databases. The database automatically generates a consensus sequence for each domain.

[0363] A BLAST search was performed against the HMM database resulting in the identification of regions homologous to ProDom family PD003768, ("nucleoside transporter cotransporter transmembrane concentrative sodium/nucleoside na/nucleoside sodium-coupled" SEQ ID NOs:64 and 65, ProDomain Release 2001.1). The "nucleoside transporter cotransporter transmembrane concentrative sodium/nucleoside na/nucleoside sodium-coupled" domain (amino acids 262-585 and 201-261 of SEQ ID NO:61) of human 52991 aligns with consensus amino acid sequences (SEQ ID NOs: 64 and 65) derived from a hidden Markov model. The consensus sequences for SEQ ID NOs:64 and 65 are 53% and 34% identical over amino acids 262 to 585 and 201 to 261 of SEQ ID NO:61, respectively.

[0364] A BLAST search was performed against the HMM database resulting in the identification of regions homologous to ProDom family PD351462, ("concentrative cotransporter Na+ nucleoside HCNT3 MCNT3" SEQ ID NO:66, ProDomain Release 2001.1. The "concentrative cotransporter na+ nucleoside HCNT3 MCNT3" domain (amino acids 587-663 of SEQ ID NO:61) of human 52991 aligns with consensus amino acid sequences (SEQ ID NO:66) derived from a hidden Markov model. The consensus sequence for SEQ ID NO:66 is 81% identical over amino acids 587 to 663 of SEQ ID NO:61.

[0365] A BLAST search was performed against the HMM database resulting in the identification of regions homologous to ProDom family PD008773, ("nucleoside cotransporter concentrative sodium/nucleoside na/nucleoside sodium-coupled transmembrane" SEQ ID NO:67, ProDomain Release 2001.1). The "nucleoside cotransporter concentrative sodium/nucleoside na/nucleoside sodium-coupled transmembrane" domain (amino acids 93-195 of SEQ ID NO:61) of human 52991 aligns with a consensus amino acid sequence (SEQ ID NO:67) derived from a hidden Markov model. The consensus sequence for SEQ ID NO:67 is 52% identical over amino acids 93 to 195 of SEQ ID NO:61.

[0366] A BLAST search was performed against the HMM database resulting in the identification of regions homologous to ProDom family PD353176 ("concentrative cotransporter Na+ nucleoside HCNT3 MCNT3" SEQ ID NO:68, ProDomain Release 2001.1). The "concentrative cotransporter Na+ nucleoside HCNT3 MCNT3" domain (amino acids 1-91 of SEQ ID NO:61) of human 52991 aligns with consensus amino acid sequences (SEQ ID NO:68) derived from a hidden Markov model. The consensus sequence for SEQ ID NO:68 is 61% identical over amino acids 1 to 91 of SEQ ID NO:61.

[0367] In one embodiment, a 52991 protein includes at least one transmembrane domain. As used herein, the term "transmembrane domain" includes an amino acid sequence of about 15 amino acid residues in length that spans a phospholipid membrane. More preferably, a transmembrane domain includes about at least 18, 20, 22, 24, 25, 30, 35 or 40 amino acid residues and spans a phospholipid membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an α-helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, e.g., leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains are described in, for example, Zagotta W. N. et al., (1996) *Annual Rev. Neuronsci.* 19: 235-63, the contents of which are incorporated herein by reference.

[0368] In a preferred embodiment, a 52991 polypeptide or protein has at least one transmembrane domain or a region which includes at least 18, 20, 22, 24, 25, 30, 35 or 40 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "transmembrane domain," e.g., at least one transmembrane domain of human 52991 (e.g., amino acid residues 104-120 of SEQ ID NO:61).

[0369] In another embodiment, a 52991 protein includes at least one "non-transmembrane domain." As used herein, "non-transmembrane domains" are domains that reside out-

side of the membrane. When referring to plasma membranes, non-transmembrane domains include extracellular domains (i.e., outside of the cell) and intracellular domains (i.e., within the cell). When referring to membrane-bound proteins found in intracellular organelles (e.g., mitochondria, endoplasmic reticulum, peroxisomes and microsomes), non-transmembrane domains include those domains of the protein that reside in the cytosol (i.e., the cytoplasm), the lumen of the organelle, or the matrix or the intermembrane space (the latter two relate specifically to mitochondria organelles). The C-terminal amino acid residue of a non-transmembrane domain is adjacent to an N-terminal amino acid residue of a transmembrane domain in a naturally-occurring 52991, or 52991-like protein.

[0370] In a preferred embodiment, a 52991 polypeptide or protein has a "non-transmembrane domain" or a region which includes at least about 1-105, preferably about 2-103, more preferably about 3-103, and even more preferably about 5-103 amino acid residues, and has at least about 60%, 70% 80% 90% 95%, 99% or 100% homology with a "non-transmembrane domain", e.g., a non-transmembrane domain of human 52991 (e.g., residues 1-103 and 587-665 of SEQ ID NO:61). Preferably, a non-transmembrane domain is capable of catalytic activity (e.g., transport of molecules across a lipid bilayer).

[0371] A non-transmembrane domain located at the N-terminus of a 52991 protein or polypeptide is referred to herein as an "N-terminal non-transmembrane domain." As used herein, an "N-terminal non-transmembrane domain" includes an amino acid sequence having about 1-105, preferably about 40-103, more preferably about 80-103, or even more preferably about 90-103 amino acid residues in length and is located outside the boundaries of a membrane. For example, an N-terminal non-transmembrane domain is located at about amino acid residues 1-103 of SEQ ID NO:61.

[0372] Similarly, a non-transmembrane domain located at the C-terminus of a 52991 protein or polypeptide is referred to herein as a "C-terminal non-transmembrane domain." As used herein, a "C-terminal non-transmembrane domain" includes an amino acid sequence having about 1-80, preferably about 30-80, preferably about 40-78, more preferably about 60-78 amino acid residues in length and is located outside the boundaries of a membrane. For example, a C-terminal non-transmembrane domain is located at about amino acid residues 587-665 of SEQ ID NO:61.

[0373] In a preferred embodiment, a 52991 family member can include at least one Na+ dependent nucleoside transporter family domain (PFAM Accession Number PF01773). Furthermore, a 52991 family member can include at least one, two, three, four, and preferably five N-glycosylation site (PS00001); at least one, two, three, four, five, six, seven, eight, and preferably nine protein kinase C phosphorylation sites (PS00005); at least one, two, three, and preferably four casein kinase II phosphorylation sites (PS00006); at least one tyrosine kinase phosphorylation site (PS00007); at least one, two, three, four, five, six, seven, eight, nine, and preferably ten N-myristolyation sites (PS00008); at least one amidation site (PS00009); and at least one, preferably two prokaryotic membrane lipoprotein lipid attachment sites (PS00013).

[0374] As the 52991 polypeptides of the invention may modulate 52991-mediated activities, they may be useful for developing novel diagnostic and therapeutic agents for 52991-mediated or related disorders, as described below.

[0375] As used herein, a "transporter-associated disorder" includes a disorder, disease or condition which is caused or characterized by a misregulation (e.g., downregulation or upregulation) of a transporter-mediated activity. Transporter-associated disorders can detrimentally affect cellular functions such as cellular proliferation, growth, differentiation, or migration, cellular regulation of homeostasis, inter- or intracellular communication; tissue function, such as cardiac function or musculoskeletal function; systemic responses in an organism, such as nervous system responses, hormonal responses (e.g., insulin response), or immune responses; and protection of cells from toxic compounds (e.g., carcinogens, toxins, mutagens, and toxic byproducts of metabolic activity (e.g., reactive oxygen species)). Accordingly, 52991 protein may mediate various disorders, including cellular proliferative and/or differentiative disorders, immune disorders, blood vessel disorders, bone metabolism, liver disorders, and pain or metabolism disorders. As the 52991 polypeptides of the invention may modulate 52991-mediated activities, they may be useful for developing novel diagnostic and therapeutic agents for 52991-mediated or related disorders, as described below.

[0376] The 52991 nucleic acid and protein of the invention can be used to treat and/or diagnose a variety of proliferative disorders, e.g., such disorders include hematopoietic neoplastic disorders.

[0377] Transporter-associated or related disorders also include immune disorders, such as autoimmune disorders or immune deficiency disorders, e.g., congenital X-linked infantile hypogammaglobulinemia, transient hypogammaglobulinemia, common variable immunodeficiency, selective IgA deficiency, chronic mucocutaneous candidiasis, or severe combined immunodeficiency.

[0378] Aberrant expression and/or activity of 52991 molecules may mediate disorders associated with bone metabolism.

[0379] Disorders which may be treated or diagnosed by methods described herein include, but are not limited to, disorders associated with an accumulation in the liver of fibrous tissue, such as that resulting from an imbalance between production and degradation of the extracellular matrix accompanied by the collapse and condensation of preexisting fibers. Additionally, 52991 may play an important role in the regulation of metabolism or pain disorders.

Gene Expression Analysis of 52991

[0380] TaqMan real-time quantitative RT-PCR is used to detect the presence of RNA transcript corresponding to human 52991 relative to a no template control in a panel of human tissues or cells. It is found that the highest expression of 52991 orthologs are expressed in osteoclasts and pancreas tissue, as shown in Table 14. Relatively high expression is also seen in normal breast tissue, ovary tumor tissue, colon tumor tissue, lung tumor tissue, and neutrophils. It is also of note that there is decreased expression of 52991 in breast tumor, and pancreatic tumor compared to normal breast and pancreatic tissue. Furthermore, there is increased expression of 52991 in ovary tumor, colon tumor, lung tumor, fibrotic liver tissue and diseased aorta tissue compared to normal ovary, colon, lung and liver tissue and normal artery tissue.

TABLE 14

Tissue Type	Mean	β 2 Mean	∂∂ Ct	Expression
Artery normal	38.56	23.51	13.75	0
Aorta diseased	34.45	23.72	9.43	1.4497
Vein normal	38.4	21.66	15.43	0
Coronary SMC	35.59	22.23	12.06	0
HUVEC	40	22.77	15.93	0
Hemangioma	38.57	21.02	16.25	0
Heart normal	36.29	21.9	13.09	0
Heart CHF	37.17	22.17	13.7	0
Kidney	33.94	21.73	10.91	0.5197
Skeletal Muscle	38.99	23.35	14.34	0
Adipose normal	36.91	22.07	13.53	0
Pancreas	30.75	23.59	5.85	17.337
Primary osteoblasts	33.66	22.11	10.24	0.8269
Osteoclasts (diff)	25.91	18.95	5.65	19.915
Skin normal	33.06	23.68	8.07	3.7212
Spinal cord normal	37.02	22.2	13.51	0
Brain Cortex normal	39.39	23.39	14.7	0
Brain Hypothalamus normal	35.68	23.86	10.52	0
Nerve	38.91	23.24	14.37	0
DRG (Dorsal Root Ganglion)	40	23.27	15.43	0
Breast normal	31.11	22.46	7.34	6.1508
Breast tumor	32.51	22.27	8.94	2.0361
Ovary normal	33.22	21.23	10.68	0.6095
Ovary Tumor	29.16	20.57	7.29	6.412
Prostate Normal	34.99	20.88	12.81	0.1393
Prostate Tumor	36.73	21.61	13.82	0
Salivary glands	33.48	21.18	10.99	0.4917
Colon normal	30.81	19.93	9.57	1.3111
Colon Tumor	31.1	22.77	7.03	7.6517
Lung normal	33.22	19.54	12.38	0.1883
Lung tumor	30.05	21.51	7.23	6.6612
Lung COPD	31.3	19.79	10.21	0.8443
Colon IBD	29.48	19.06	9.12	1.8035
Liver normal	34.72	21.63	11.79	0.2834
Liver fibrosis	32.23	22.04	8.88	2.1152
Spleen normal	36.28	19.91	15.06	0
Tonsil normal	30.25	19.33	9.62	1.2708
Lymph node normal	33.69	21.07	11.31	0.3939
Small intestine normal	32.12	21.54	9.28	1.6086
Macrophages	33.07	18.34	13.43	0.0906
Synovium	32.27	21.02	9.95	1.011
BM-MNC	31.01	20.34	9.37	1.5165
Activated PBMC	36.84	19.38	16.16	0
Neutrophils	28.73	20.36	7.06	7.4943
Megakaryocytes	38.08	20.26	16.51	0
Erythroid	39.47	23.13	15.03	0
positive control	32.78	21.49	9.99	0.9834

Human 59914 and 59921

[0381] The present invention is based, in part, on the discovery of novel genes encoding choline transporters, the genes being referred to herein as "59914 and 59921".

[0382] The human 59914 cDNA sequence (SEQ ID NO:69), which is approximately 2473 nucleotide residues long including non-translated regions, contains a methionine-initiated coding sequence (without the 5'- and 3'-non-translated regions) of about 2151 nucleotide residues, excluding termination codon (i.e., nucleotide residues 88-2238 of SEQ ID NO:69; 1-2151 of SEQ ID NO:71). The coding sequence encodes a 717 amino acid protein having the amino acid sequence SEQ ID NO:70.

[0383] The human 59921 cDNA sequence (SEQ ID NO:72), which is approximately 2233 nucleotide residues long including non-translated regions, contains a methionine-initiated coding sequence (without the 5'- and 3'-non-translated regions) of about 1959 nucleotide residues, excluding termination codon (i.e., nucleotide residues 110-2068 of SEQ ID NO:72; 1-1959 of SEQ ID NO:74). The coding sequence encodes a 653 amino acid protein having the amino acid sequence SEQ ID NO:73.

[0384] Human 59914 and 59921 contain the following regions or other structural features: 1) a conserved region of sequence which is shared by both 59914 and 59921 proteins, and by other choline transporter (or choline transporter-like) proteins described herein and in O'Regan et al. (2000) *PNAS* 97(4):1835-1840. This region will henceforth be referred to as "conserved choline transporter domain", and is located at about amino acid residues 479-598 of SEQ ID NO:70 and about amino acid residues 402-521 of SEQ ID NO:73; 2) transmembrane domains at about amino acid residues 39-61, 242-263, 270-287, 326-346, 371-395, 461-484, 514-536, 591-605, 608-632, and 649-672 of SEQ ID NO:70, and at about 33-57, 215-231, 239-262, 284-305, 328-352, 384-411, 436-458, 514-532, 534-555, and 563-586 of SEQ ID NO:73, 59914 and 59921 proteins therefore have about 10 transmembrane domains, as is characteristic of previously characterized choline transporters; 3) conserved cysteine residues at about amino acid residues 36, 79, 117, 121, 149, 168, 182, 557, 558, 561, and 681 of SEQ ID NO:70 and about amino acid residues 29, 72, 116, 120, 143, 162, 179, 480, 481, 484, and 596 of SEQ ID NO:73, and also found in other choline transporter (and choline transporter-like) proteins described herein and in O'Regan et al. (2000) *PNAS* 97(4):1835-1840. 59914 and 59921 proteins therefore have about 11 conserved cysteines, as is characteristic of previously characterized choline transporters; 4) and post translational modification sites including: predicted N-glycosylation sites (Pfam accession number PS00001) at about amino acid residues 2-5, 33-36, 88-91, 190-193, 314-317, 416-419, and 425-428 of SEQ ID NO:70 and at about amino acid residues 136-139, 151-154, 412-415, 503-506, and 521-524 of SEQ ID NO:73; predicted protein kinase C phosphorylation sites (Pfam accession number PS00005) at about amino acid residues 4-6, 152-154, 187-189, 587-589, 633-635, 693-695, and 714-716 of SEQ ID NO:70 and at about amino acid residues 90-92, 155-157, 210-212, 232-234, 276-278, 319-321, 510-512, 608-610, 625-627, and 639-641 of SEQ ID NO:73; predicted cAMP- and cGMP-dependent kinase phosphorylation sites (Pfam accession number PS00004) at about amino acid residues 27-30 and 74-77 of SEQ ID NO:73; predicted casein kinase II phosphorylation sites (Pfam accession number PS00006) located at about amino acid residues 10-13, 35-38, 84-87, 127-130, 140-143, 210-213, 305-308, and 587-590 of SEQ ID NO:70 and at about amino acid residues 77-80, 85-88, 127-130, 210-213, 276-279, 414-417, 510-513, 584-587, and 588-591 of SEQ ID NO:73; predicted N-myristoylation sites (Pfam accession number PS00008) at about amino acid residues 204-209, 215-220, 248-253, 285-290, 310-315, 435-440, and 481-486 of SEQ ID NO:70 and at about amino acid residues 69-74, 81-86, 107-112, 139-144, 207-212, 355-360, 386-391, 404-409, 504-509, and 550-555 of SEQ ID NO:73; a predicted tyrosine kinase phosphorylation site (Pfam accession number PS00007) at about amino acid residues 491-499 of SEQ ID NO:70; and a predicted amidation site (Pfam accession number PS00009) at about amino acid residues 72-75 of SEQ ID NO:73.

[0385] A hydropathy plot of human 59914 was performed. Polypeptides of the invention include fragments which include: all or part of a hydrophobic sequence, e.g., the sequences of about residues 39-61, 242-263, and 326-346 of SEQ ID NO:70; all or part of a hydrophilic sequence, e.g., the

sequences of residues 62-90, 347-370, and 633-648 of SEQ ID NO:70; a sequence which includes a cysteine residue; or a glycosylation site.

[0386] A hydropathy plot of human 59921 was performed. Polypeptides of the invention include fragments which include: all or part of a hydrophobic sequence, e.g., the sequences of about residues 33-57, 215-231, and 328-352 of SEQ ID NO:73; all or part of a hydrophilic sequence, e.g., the sequences of residues 58-214, 263-283, and 306-327 of SEQ ID NO:73; a sequence which includes a cysteine residue; or a glycosylation site.

[0387] The 59914 and 59921 proteins contain a number of structural characteristics in common with members of the choline transporter family. Choline transporter family members all show several transmembrane domains and can reasonably be thought to traverse the membrane about 10 times. In one embodiment, a first, large and variable loop between transmembrane domains 1 and 2 is potentially extracellular and glycosylated. In one embodiment, a highly conserved region covers the last four transmembrane domains and includes the fourth extracellular loop that contains about three conserved cysteines. Choline transporter family members generally lack a clear signal peptide and are targeted to the plasma membrane via their transmembrane domains.

[0388] 59914 and 59921 proteins can include a conserved choline transporter domain. As used herein, a "conserved choline transporter domain" refers to a protein domain having an amino acid sequence of about 50-250 amino acid residues in length, preferably about 75-175 amino acid residues in length, more preferably about 100-150 amino acid residues in length, and most preferably about 119-121 amino acid residues in length; and which has about 1-10 conserved cysteine residues, preferably about 2-8 conserved cysteine residues, and more preferably about 3-7 conserved cysteine residues.

[0389] In one embodiment, the conserved choline transporter domain can have one, preferably both, of the following consensus sequences: [LVI]-A-G-A-Xaa(2)-[ST]-[CY]-Y-[FW]-Xaa(3)-K-Xaa(n1)-P-Xaa(2)-P-[LI]-Xaa(5)-[IR]-Xaa(3)-Y-H-Xaa-G-Xaa(4)-G-Xaa(2)-[LI]-[LI]-Xaa(4)-[IM]-Xaa(2)-[VMI]-[VI]-[VL] (SEQ ID NO:79) and/or L-K-[ERG]-Xaa(2)-[HN]-Xaa(n2)-C-C-Xaa-W-C-L-[DE]-Xaa(8)-N-A-Y-Xaa(3)-[AS]-I-Xaa(4)-F-C-Xaa-S-A-K-D-A-[FI]-Xaa-[IL]-L-Xaa(2)-N (SEQ ID NO:80)

[0390] In these consensus sequence patterns, each element in the pattern is separated by a dash (-); square [ ] brackets indicate the particular residues that are accepted at that position; Xaa indicates any residue is accepted at that position; repetition of a particular element is indicated by following the element with a numerical value or variable enclosed in parentheses (i.e., above, Xaa(2) indicates 2 residues of any type are repeated, and Xaa(n1) indicates that a range of residues of any type are repeated, as described herein); and the standard IUPAC one-letter code for the amino acids is used. n1 in the first consensus sequence (SEQ ID NO:79) can be 1-8, preferably 2-6, more preferably 3-4, and n2 in the second consensus sequence (SEQ ID NO:80) can be 8-15, preferably 10-13, more preferably 11-12.

[0391] These consensus sequences are found from about residues 479-533 and 540-598, of the 59914 protein (of SEQ ID NO:70), respectively; and from about residues 402-455 and 462-521 of the 59921 protein (of SEQ ID NO:73), respectively.

[0392] A conserved choline transporter domain is found in at least the following choline transporter (or choline transporter-like (CTL)) proteins: human CTL1 (Genbank accession number CAB75541; SEQ ID NO:75); human CTL2 (Genbank accession number CAB75542; SEQ ID NO:76); rat CTL1 (Genbank accession number CAB75555; SEQ ID NO:77); and torpedo CTL1 (Genbank accession number CAB75556; SEQ ID NO:78). For example, in the human CTL1 protein (Genbank accession number CAB75541; SEQ ID NO:75), the conserved choline transporter domain as described herein is found from about amino acids 407-525 (of SEQ ID NO:75), and the consensus sequences as described herein (SEQ ID NO:79 and SEQ ID NO:80) are found from about amino acids 407-460 and 467-525, respectively (of SEQ ID NO:75).

[0393] In one embodiment, the 59914 and 59921 polypeptides or proteins have a conserved choline transporter domain which includes at least about 50-250, preferably about 75-175, more preferably about 100-150, and most preferably about 119-121 amino acid residues in length and has at least about 60%, 70%, 80%, 90%, 95%, 99%, or 100% homology with a conserved choline transporter domain, e.g., the conserved choline transporter domain of human 59914 or 59921 (e.g., residues 479-598 of SEQ ID NO:70 or residues 402-521 of SEQ ID NO:73). In another embodiment, the 59914 and 59921 polypeptides or proteins have a conserved choline transporter domain which includes at least about 50-250, preferably about 75-175, more preferably about 100-150, and most preferably about 119-121 amino acid residues in length; has at least about 60%, 70%, 80%, 90%, 95%, 99%, or 100% homology with a conserved choline transporter domain, e.g., the conserved choline transporter domain of human 59914 or 59921 (e.g., residues 479-598 of SEQ ID NO:70 or residues 402-521 of SEQ ID NO:73); and has about 1-10, preferably about 2-8, and more preferably about 3-7 conserved cysteine residues (e.g., at about positions 553, 554, 557, 558, 561, and 585 of SEQ ID NO:70, and at about positions 409, 477, 478, 480, 481, 484, and 508 of SEQ ID NO:73).

[0394] In another embodiment, the 59914 and 59921 polypeptides or proteins have a conserved choline transporter domain which includes at least about 50-250, preferably about 75-175, more preferably about 100-150, and most preferably about 119-121 amino acid residues in length; has at least about 60%, 70%, 80%, 90%, 95%, 99%, or 100% homology with a conserved choline transporter domain, e.g., the conserved choline transporter domain of human 59914 or 59921 (e.g., residues 479-598 of SEQ ID NO:70 or residues 402-521 of SEQ ID NO:73); has about 1-10, preferably about 2-8, and more preferably about 3-7 conserved cysteine residues (e.g., at about positions 553, 554, 557, 558, 561, and 585 of SEQ ID NO:70, and at about positions 409, 477, 478, 480, 481, 484, and 508 of SEQ ID NO:73); and has one or more of the conserved choline transporter domain consensus sequences described herein. In still another embodiment, the 59914 and 59921 polypeptides or proteins have a conserved choline transporter domain which includes at least about 50-250, preferably about 75-175, more preferably about 100-150, and most preferably about 119-121 amino acid residues in length; has at least about 60%, 70%, 80%, 90%, 95%, 99%, or 100% homology with a conserved choline transporter domain, e.g., the conserved choline transporter domain of human 59914 or 59921 (e.g., residues 479-598 of SEQ ID NO:70 or residues 402-521 of SEQ ID NO:73); has about 1-10, preferably about 2-8, and more preferably about 3-7 conserved cysteine residues (e.g., at about positions 553, 554, 557, 558, 561, and 585 of SEQ ID NO:70, and at about

positions 409, 477, 478, 480, 481, 484, and 508 of SEQ ID NO:73); has one or more of the conserved choline transporter domain consensus sequences described herein; and has at least one 59914 and 59921 biological activity as described herein

[0395] In one embodiment, 59914 and 59921 proteins include at least ten transmembrane domains. As used herein, the term "transmembrane domain" includes an amino acid sequence of about 5 amino acid residues in length that spans the plasma membrane. More preferably, a transmembrane domain includes about at least 10, 15, 20 or 22-25 amino acid residues and spans a membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an alpha-helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, or 95% or more of the amino acids of a transmembrane domain are hydrophobic, e.g., leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains are described in, for example, Zagotta W. N. et al. (1996) *Annu. Rev. Neurosci.* 19:235-263, the contents of which are incorporated herein by reference. Transmembrane domains exist at least from about amino acid residues 39-61, 242-263, 270-287, 326-346, 371-395, 461-484, 514-536, 591-605, 608-632, and 649-672 of SEQ ID NO:70, and at least from about amino acid residues 33-57, 215-231, 239-262, 284-305, 328-352, 384-411, 436-458, 514-532, 534-555, and 563-586 of SEQ ID NO:73.

[0396] A 59914 and 59921 family member can include at least one conserved choline transporter domain. Furthermore, a 59914 and 59921 family member can include at least one, preferably at least 5, more preferably at least 9, and still more preferably 10 transmembrane domains; at least one, preferably 5-7, N-glycosylation sites; at least one, preferably 7-10, protein kinase C phosphorylation sites; at least one, preferably 8-9 casein kinase II phosphorylation sites; and at least one, preferably 7-10 N-myristoylation sites.

[0397] 59914 and 59921 are homologous to human CTL1 (Genbank accession number CAB75541; SEQ ID NO:75) and human CTL2 (Genbank accession number CAB75542; SEQ ID NO:76), both human choline transporter-like proteins known in the art. An alignment of hCTL1 with the amino acid sequence of 59914 (SEQ ID NO:70) reveals 39.7% identity and 29.4% homology. An alignment of hCTL2 with the amino acid sequence of 59914 (SEQ ID NO:70) reveals 66.6% identity and 54.8% homology. An alignment of hCTL1 with the amino acid sequence of 59921 (SEQ ID NO:73) reveals 57.3% identity and 47.7% homology. An alignment of hCTL2 with the amino acid sequence of 59921 (SEQ ID NO:73) reveals 41.5% identity and 29.6% homology. [The alignments described in this paragraph were performed using the GAP alignment program with a BLOSUM62 scoring matrix, a gap open penalty of 12, and a gap extend penalty of 4.]

[0398] Like the 59914 and 59921 proteins, hCTL1 and hCTL2 contain a conserved choline transporter domain (from about amino acid residues 407-525 of SEQ ID NO:75 and about amino acid residues 468-587 of SEQ ID NO:76, respectively); 10 transmembrane domains, as described in O'Regan, supra; and 10 conserved cysteines, as described in O'Regan, supra.

[0399] 59914 and 59921 are also homologous to rat and torpedo (marbled electric ray) choline transporter like proteins (rCTL1 (Genbank accession number CAB75555; SEQ ID NO:77) and tCTL1 (Genbank accession number CAB75556; SEQ ID NO:78), respectively). rCTL1 and tCTL1 are described in O'Regan, supra, and were discovered in the context of suppressing a yeast choline transport mutation (the addition of tCTL1 to yeast increased high-affinity choline uptake in mutant yeast). An alignment of rCTL1 with the amino acid sequence of 59914 (SEQ ID NO:70) reveals 39.5% identity and 29.2% homology. An alignment of tCTL1 with the amino acid sequence of 59914 (SEQ ID NO:70) reveals 40.2% identity and 29.3% homology. An alignment of rCTL1 with the amino acid sequence of 59921 (SEQ ID NO:73) reveals 57.5% identity and 47.4% homology. An alignment of tCTL1 with the amino acid sequence of 59921 (SEQ ID NO:73) reveals 56.3% identity and 45.2% homology. [The alignments described in this paragraph were performed using the GAP alignment program with a BLOSUM62 scoring matrix, a gap open penalty of 12, and a gap extend penalty of 4.]

[0400] Like the 59914 and 59921 proteins, rCTL1 and tCTL1 contain a conserved choline transporter domain (from about amino acid residues 406-524 of SEQ ID NO:77 and about amino acid residues 399-518 of SEQ ID NO:78, respectively); 10 transmembrane domains, as described in O'Regan, supra; and 10 conserved cysteines, as described in O'Regan, supra.

[0401] Based on the above described sequence similarities, the 59914 and 59921 molecules of the present invention belong to the choline transporter family (as described herein). Consequently, the 59914 and 59921 molecules of the invention have similar biological activities as choline transporter family members, and are useful in treating the same disorders as choline transporter family members.

[0402] Based on sequence similarities of 59914 and 59921 to sequences of known expression pattern, 59914 and 59921 molecules of the invention can exhibit similar expression patterns, and therefore can be useful in treating disorders associated with tissues in which they are expressed.

[0403] Because the 59914 and 59921 polypeptides of the invention can modulate 59914 and 59921-mediated activities, they can be used as novel diagnostic and therapeutic agents or used to develop novel diagnostic and therapeutic agents for 59914 and 59921-mediated or related disorders (e.g., disorders associated with choline transporter family members), as described below.

[0404] As used herein, a "59914 and 59921 activity", "biological activity of 59914 and 59921", or "functional activity of 59914 and 59921", refers to an activity of a choline transporter family member, and refers to an activity exerted by 59914 and 59921 proteins, polypeptides or nucleic acid molecules on, for example, 59914 and 59921-responsive cells or on 59914 and 59921 substrates (e.g., protein substrates) as determined in vivo or in vitro. In one embodiment, a 59914 and 59921 activity is a direct activity, such as association with 59914 and 59921 target molecules. "Target molecules" or "binding partners" of 59914 and 59921 proteins are molecules with which the 59914 and 59921 proteins bind or interact in nature. In an exemplary embodiment, such target molecules include choline, its metabolites, and/or compounds of which choline is a component or precursor, e.g., which 59914 and 59921 proteins can transport into cells from the extracellular fluid, e.g., for plasma membrane synthesis.

[0405] A 59914 and 59921 activity can also be an indirect activity, such as an activity mediated by interaction of the 59914 and 59921 protein with a 59914 and 59921 target molecule such that the target molecule modulates a downstream cellular activity, e.g., a cellular signaling activity

modulated indirectly by interaction of the 59914 and 59921 protein with a 59914 and 59921 target molecule (e.g., choline, its metabolites, and/or compounds of which choline is a component or precursor).

[0406] For example, the 59914 and 59921 proteins of the present invention can have one or more of the following activities: (1) the ability to modulate (e.g., promote, catalyze, regulate, initiate, facilitate or inhibit) the manufacture of choline metabolites and/or compounds of which choline is a component or precursor, e.g., phospholipids (e.g., phosphatidylcholine (lecithin), sphingomyelin, sphingophosphorylcholine, and platelet activating factor), acetylcholine, very low density lipoproteins (VLDLs), and betaine, e.g., by transporting choline into or out of cells; (2) the ability to modulate (e.g., promote, catalyze, regulate, initiate, facilitate or inhibit) transport of choline, its metabolites, and/or compounds of which choline is a component or precursor across membranes (e.g., plasma membranes), e.g., from an extracellular medium into a cell, or vice versa; (3) the ability to modulate (e.g., promote, catalyze, regulate, initiate, facilitate or inhibit) transport of choline, its metabolites, and/or compounds of which choline is a component or precursor across barriers between tissues (e.g., the blood-brain barrier).

[0407] Other activities of the 59914 and 59921 proteins of the present invention include one or more of the following: (1) the ability to modulate (e.g., promote, regulate, initiate, facilitate or inhibit) the synthesis of, and the structural maintenance and reinforcement of, cellular components (e.g., membranes (e.g., plasma membranes) and microsomes); (2) the ability to modulate (e.g., promote, regulate, initiate, facilitate or inhibit) cellular nutrition; (3) the ability to modulate (e.g., promote, regulate, initiate, facilitate or inhibit) muscle control; (4) the ability to modulate (e.g., promote, regulate, initiate, facilitate or inhibit) memory; and (5) the ability to modulate (e.g., promote, regulate, initiate, facilitate or inhibit) message transmission (e.g., nervous system message transmission).

[0408] Still other activities of the 59914 and 59921 proteins of the present invention include one or more of the following: (1) the ability to modulate (e.g., promote, regulate, initiate, facilitate or inhibit) liver homeostasis, e.g., by transporting fat and/or cholesterol from the liver, e.g., by modulating the transport of choline, its metabolites, and/or compounds of which choline is a component or precursor (e.g., VLDLs); and (2) the ability to modulate (e.g., promote, regulate, initiate, facilitate or inhibit) cellular signaling, e.g., by modulating the transport of choline, its metabolites, and/or compounds of which choline is a component or precursor (e.g., sphingophosphorylcholine and platelet activating factor).

[0409] Other activities of the 59914 and 59921 proteins of the present invention include one or more of the following: (1) the ability to modulate (e.g., promote, regulate, initiate, facilitate or inhibit) liver disorders (e.g., hepatocyte apoptosis and others described herein), e.g., by maintaining proper choline levels (e.g., by preventing choline deficiency), e.g., by modulating the transport of choline, its metabolites, and/or compounds of which choline is a component or precursor; (2) the ability to modulate (e.g., promote, regulate, initiate, facilitate or inhibit) central nervous system (CNS) disorders (e.g., hepatocyte apoptosis, and others described herein), e.g., by maintaining proper choline levels (e.g., by preventing choline excess), e.g., by modulating the transport of choline, its metabolites, and/or compounds of which choline is a component or precursor; and (3) the ability to modulate (e.g., promote, regulate, initiate, facilitate or inhibit) cardiovascular disorders, e.g., by preventing buildup of homocysteines in the blood (e.g., by converting them to methionine) e.g., by modulating the transport of choline, its metabolites, and/or compounds of which choline is a component or precursor (e.g., betaine).

[0410] Other activities, as described below, include the ability to modulate function, survival, morphology, proliferation and/or differentiation of cells of tissues in which 59914 and 59921 molecules are expressed. Thus, the 59914 and 59921 molecules can act as novel diagnostic targets and therapeutic agents for controlling disorders involving aberrant activities of these cells. 59914 and 59921 molecules described herein can act as novel diagnostic targets and therapeutic agents for prognosticating, diagnosing, preventing, inhibiting, alleviating, or curing choline transporter-related disorders. As the 59914 and 59921 molecules of the invention can modulate choline transporter activities, they are useful for developing novel diagnostic and therapeutic agents for 59914 and 59921-mediated or related disorders, as described herein.

[0411] As used herein, a "choline transporter disorder" includes a disorder, disease or condition which is caused by, characterized by, or associated with a misregulation (e.g., an aberrant downregulation or upregulation) of an choline transporter activity or an abnormal choline transporter activity. Choline transporter disorders can detrimentally affect cellular functions such as amino acid nutrition, cellular regulation of homeostasis, membrane structural integrity, and inter- or intra-cellular communication.

[0412] Accordingly, the 59914 and 59921 molecules of the invention, as choline transporters, can mediate, and can act as novel diagnostic targets and therapeutic agents for controlling, one or more choline transporter-associated disorders, including CNS-related (e.g., neurological) disorders; liver-related (i.e., hepatic) disorders; skeletal muscle-related disorders; lung-related (i.e., pulmonary) disorders, prostate-related disorders, kidney-related (i.e., renal) disorders, pancreas-related disorders, colon-related disorders, cellular proliferative and/or differentiative disorders; hormonal disorders; immune and inflammatory disorders; cardiovascular disorders; blood vessel disorders; and platelet disorders.

[0413] Further, polymorphisms associated with particular 59914 and 59921 alleles, such as those associated with risk of choline transporter-associated disorders, can be used as markers to diagnose abnormal function of tissues and/or cells in which 59914 and 59921 are expressed (described herein), and therefore can be used as markers for disorders associated with such tissues. For example, abnormal and/or aberrant 59914 and 59921 expression (e.g., expression of 59914 and 59921 in cells, such as tumor cells, that do not normally express them, or increased expression of 59914 and 59921 in cells that do normally express them) can be used as a marker for the progression, migration and metastasis of cancerous cells. In particular, abnormal and/or aberrant 59914 and 59921 expression can be used as a marker for the progression, migration and metastasis of cancers of the tissues and/or cells in which 59914 and 59921 are expressed (described herein).

[0414] Additional choline transporter disorders include CNS-related (e.g., neurological) disorders, hepatic disorders, skeletal muscle-related disorders, pulmonary (lung) disorders, prostate disorders, renal (kidney) disorders, pancreatic disorders and colonic disorders.

[0415] The 59914 and 59921 molecules of the invention can be used to monitor, treat and/or diagnose a variety of proliferative disorders. Such disorders include hematopoietic neoplastic disorders.

[0416] Choline transporter disorders can include hormonal disorders, such as conditions or diseases in which the production and/or regulation of hormones in an organism is aberrant.

[0417] Choline transporter disorders also include immune disorders, such as autoimmune disorders or immune deficiency disorders.

Gene Expression of 59914 and 59921

[0418] TaqMan analysis indicates significant 59914 expression in normal brain cortex; moderate 59914 expression in human umbilical vein endothelial cells (HUVEC), prostate tumor and lung tumor; low levels of 59914 expression in colon tumor, kidney, and hypothalamus. It also indicates significant 59921 expression in kidney, pancreas, and colon tumor; and low to moderate 59921 levels of expression in spinal cord, hypothalamus, nerve, dorsal root ganglia, prostate tumor, lung tumor, salivary glands, and liver fibrosis

Human 33751

[0419] Nucleotide and corresponding amino acid sequences for an ion channel family member, referred to herein as "33751" are disclosed. 33751 protein is a member of a family of voltage-gated potassium channel genes that includes the eag, erg, and elk genes.

[0420] The human 33751 sequence (SEQ ID NO:81), which is approximately 4113 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 3588 nucleotides long (nucleotides 101-3688 of SEQ ID NO:81; 1-3588 of SEQ ID NO:83), not including the termination codon. The coding sequence encodes a 1196 amino acid protein (see SEQ ID NO:82).

[0421] Human 33751 contains the following regions or other structural features: one predicted ion transport protein domain (PFAM Accession Number PF00520) located at about amino acid residues 450 to 662 of SEQ ID NO:82; one predicted PAS domain (PFAM Accession Number PF00989) located at about amino acids 41 to 60 of SEQ ID NO:82; one predicted PAC domain (PFAM Accession Number PF00785) located at about amino acids 93 to 120 of SEQ ID NO:82; one predicted cyclic nucleotide-binding domain (PFAM Accession Number PF00027) located at about amino acids 760 to 850 of SEQ ID NO:82; six predicted transmembrane segments located at about amino acids 412 to 433, 453 to 470, 495 to 513, 549 to 573, 614 to 630, and 642 to 666 of SEQ ID NO:82; one predicted N-terminal cytoplasmic domain located at about amino acids 1 to 411 of SEQ ID NO:82; one predicted C-terminal cytoplasmic domain located at about amino acids 667 to 1196 of SEQ ID NO:82; two predicted cytoplasmic loops located at about amino acids 471 to 494, and 574 to 613 of SEQ ID NO:82; three predicted extracellular loops located at about amino acids 434 to 452, 514 to 548, and 631 to 641 of SEQ ID NO:82; ten predicted N-glycosylation sites (PS00001) located at about amino acids 102 to 105, 230 to 233, 338 to 341, 369 to 372, 600 to 603, 661 to 664, 736 to 739, 881 to 884, 905 to 908, and 1139 to 1142 of SEQ ID NO:82; four predicted cAMP- and cGMP-dependent protein kinase phosphorylation sites (PS00004) located at about amino acids 171 to 174, 271 to 274, 346 to 349, and 893 to 896 of SEQ ID NO:82; twenty-one predicted Protein Kinase C phosphorylation sites (PS00005) located at about amino acids 74 to 76, 169 to 171, 187 to 189, 239 to 241, 269 to 271, 344 to 346, 354 to 356, 371 to 373, 392 to 394, 528 to 530, 582 to 584, 609 to 611, 639 to 641, 673 to 675, 869 to 871, 916 to 918, 922 to 924, 974 to 976, 985 to 987, 1096 to 1098, and 1099 to 1101 of SEQ ID NO:82; twenty-four predicted Casein Kinase II phosphorylation sites (PS00006) located at about amino acids 55 to 58, 133 to 136, 209 to 212, 275 to 278, 317 to 320, 514 to 517, 609 to 612, 637 to 640, 793 to 796, 821 to 824, 829 to 832, 857 to 860, 879 to 882, 883 to 886, 899 to 902, 906 to 909, 939 to 942, 963 to 966, 985 to 988, 1020 to 1023, 1027 to 1030, 1091 to 1094, 1134 to 1137, and 1170 to 1173 of SEQ ID NO:82; three predicted Tyrosine kinase phosphorylation sites (PS00007) located at about amino acids 92 to 99, 241 to 248, and 440 to 446 of SEQ ID NO:82; ten predicted N-myristylation sites (PS00008) located at about amino acids 515 to 520, 524 to 529, 592 to 597, 660 to 665, 713 to 718, 748 to 753, 951 to 956, 958 to 963, 1018 to 1023, and 1129 to 1134 of SEQ ID NO:82; and one predicted Amidation site (PS00009) located at about amino acids 595 to 598 of SEQ ID NO:82.

[0422] For general information regarding PFAM identifiers, PS prefix and PF prefix domain identification numbers, refer to Sonnhammer et al. (1997) Protein 28:405-420.

TABLE 15

Summary of transmembrane domains.

Transmembrane segment located in SEQ ID NO: 82	Cytoplasmic sequence located in SEQ ID NO: 82	Extracellular sequence located in SEQ ID NO: 82
about 412 to 433	about 1 to 411 (N-terminal)	
about 453 to 470		about 434 to 452
about 495 to 513	about 471 to 494	
about 549 to 573		about 514 to 548
about 614 to 630	about 574 to 613	
about 642 to 666		about 631 to 641
	about 667 to 1196 (C-terminal)	

[0423] A hydropathy plot of human 33751 was performed. Polypeptides of the invention include fragments which include: all or part of a hydrophobic sequence, e.g., the sequence from about amino acid 125 to 132, from about 380 to 388, and from about 782 to 790 of SEQ ID NO:82; all or part of a hydrophilic sequence, e.g., the sequence of from about amino acid 72 to 84, from about 285 to 310, and from about 880 to 902 of SEQ ID NO:82; a sequence which includes a Cys, or a glycosylation site.

[0424] The 33751 protein contains a significant number of structural characteristics in common with members of the potassium channel family. As used herein, a "potassium channel" includes a protein or polypeptide that is involved in receiving, conducting, and transmitting signals in an electrically excitable cell, e.g., a neuronal cell or a muscle cell. Potassium channels are potassium ion selective, and can determine membrane excitability (the ability of, for example, a neuron to respond to a stimulus and convert it into an impulse). Potassium channels can also influence the resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation. Potassium channels are typically expressed in electrically excitable cells, e.g., neurons, muscle, endocrine, and egg cells, and may form heteromultimeric structures, e.g., composed of pore-

forming a and cytoplasmic b subunits. Potassium channels may also be found in nonexcitable cells (e.g., thymus cells), where they may play a role in, e.g., signal transduction. Potassium channel proteins contain six transmembrane helices, wherein the last two helices flank a loop (a P-loop) which determines potassium ion selectivity. Examples of potassium channels include: (1) the voltage-gated potassium channels, (2) the ligand-gated potassium channels, e.g., neurotransmitter-gated potassium channels, and (3) cyclic-nucleotide-gated potassium channels. Voltage-gated and ligand-gated potassium channels are expressed in the brain, e.g., in brainstem monoaminergic and forebrain cholinergic neurons, where they are involved in the release of neurotransmitters, or in the dendrites of hippocampal and neocortical pyramidal cells, where they are involved in the processes of learning and memory formation. For a detailed description of potassium channels, see Kandel E. R. et al., *Principles of Neural Science*, second edition, (Elsevier Science Publishing Co., Inc., N.Y. (1985)), the contents of which are incorporated herein by reference.

[0425] 33751 polypeptides belong to a small gene family of voltage-gated potassium channels that includes the eag, erg, and elk genes (Shi, W. et al. (1997) supra). These genes have been described either *Drosophila* or mammals (Warmke J, Ganetzky B (1994) *Proc Natl Acad Sci USA* 91:3438-3442; Ludwig J. et al. (1994) *EMBO J.* 13:4451-4458; Warmke J, Drysdale R, Ganetzky B (1991) *Science* 252:1560-1562; Titus S A, Warmke J W, Ganetzky B (1997) *J Neurosci* 17:875-881). The 33751 polypeptides are highly homologous to the human erg1, and rat erg2 and erg3 previously identified (Shi, W. et al. (1997) supra). These channels share the six membrane-spanning architecture of the Kv class (Shaker-related) of voltage-gated potassium channels, but otherwise are distantly related to the Kv class channels. The channels encoded by the eag-related genes are relatively slowly activating, as compared with Kv class potassium channels (Ludwig J. et al. (1994) supra), and have some similarities to slowly activating potassium currents that are important in determining the threshold firing properties of neurons (Brown D A (1988) *M currents. In: Ion channels* (Narahashi T, ed), pp 55-94. New York: Plenum; Yamada W M et al. (1989) Multiple channels and calcium dynamics. *In: Methods in neutronal modeling* (Koch C, Segev I, eds), pp 97-133. Cambridge, Mass.: Bradford; Wang H S, McKinnon D (1996) *J Physiol* (Lond) 492:467-478). As might be anticipated from the biophysical properties of these channels, mutations in either the eag gene or the erg gene of *Drosophila* result in a hyperexcitable phenotype (Titus et al. (1997) supra; Wang et al. (1997) supra.

[0426] Accordingly, 33751-activity may be involved in neurological processes, including PLC-mediated conductances associated with the propagation of action potentials, synaptic transmission, as well nociceptive responses, and neuropathic pain, as described in more detail below.

[0427] A 33751 polypeptide can include an "ion transport protein domain", or regions homologous with an "ion transport protein domain."

[0428] As used herein, the term "ion transport protein domain" includes an amino acid sequence of about 150 to 280 amino acid residues in length and having a bit score for the alignment of the sequence to the ion transport protein domain profile (Pfam HMM) of at least 50. Preferably, an ion transport protein domain includes at least about 180 to 250 amino acids, more preferably about 200 to 220 amino acid residues, or about 213 amino acids and has a bit score for the alignment of the sequence to the ion transport protein domain (HMM) of at least 60, preferable 80, 95 or greater. The ion transport protein domain (HMM) has been assigned the PFAM Accession Number PF00520. The ion transport protein domain (amino acids 450 to 662 of SEQ ID NO:82) of human 33751 aligns with a consensus amino acid sequence (SEQ ID NO:84) derived from a hidden Markov model.

[0429] In a preferred embodiment 33751 polypeptide or protein has an "ion transport protein domain" or a region which includes at least about 150 to 280, more preferably about 180 to 250, or 200 to 220 amino acid residues and has at least about 50%, 60%, 70% 80% 90% 95%, 99%, or 100% homology with an "ion transport protein domain," e.g., the ion transport protein domain of human 33751 (e.g., residues 450 to 662 of SEQ ID NO:82).

[0430] To identify the presence of an "ion transport protein domain" in a 33751 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against the Pfam database of HMMs (e.g., the Pfam database, release 2.1) using the default parameters. For example, the hmmsf program, which is available as part of the HMMER package of search programs, is a family specific default program for MILPAT0063 and a score of 15 is the default threshold score for determining a hit. Alternatively, the threshold score for determining a hit can be lowered (e.g., to 8 bits). A description of the Pfam database can be found in Sonhammer et al. (1997) Proteins 28(3): 405-420 and a detailed description of HMMs can be found, for example, in Gribskov et al. (1990) Meth. Enzymol. 183: 146-159; Gribskov et al. (1987) Proc. Natl. Acad. Sci. USA 84: 4355-4358; Krogh et al. (1994) J. Mol. Biol. 235: 1501-1531; and Stultz et al. (1993) Protein Sci. 2: 305-314, the contents of which are incorporated herein by reference. A search was performed against the HMM database resulting in the identification of an "ion transport protein domain" in the amino acid sequence of human 33751 at about residues 450 to 662 SEQ ID NO:82.

[0431] A 33751 polypeptide can further include a "PAS domain" or regions homologous with a "PAS domain." A PAS domain appears in archaea, eubacteria and eukarya. PAS domain have been found in EAG-like K+-channels.

[0432] As used herein, a "PAS domain" includes an amino acid sequence of about 10 to 100 amino acid residues in length that is involved in ligand and/or protein-protein interactions. Preferably, the PAS domain interacts with the body of the channel, affecting gating, inactivation, and/or voltage sensitivity. Preferably, the PAS domain is located at the N-terminal cytoplasmic region of the 33751 polypeptide. The term "PAS domain" includes an amino acid sequence of about 10 to 100 amino acid residues in length and having a bit score for the alignment of the sequence to the cyclic nucleotide binding domain (HMM) of at least 5. Preferably, a PAS domain includes at least about 10 to 100 amino acids, more preferably about 13 to 50 amino acid residues, or about 17 to 25 amino acids and has a bit score for the alignment of the sequence to the PAS domain (HMM) of at least 6 or greater. The PAS domain (HMM) has been assigned the PFAM Accession PF00989. The PAS domain (amino acids 41 to 60 of SEQ ID NO:82) of human 33751 aligns with a consensus amino acid sequence (SEQ ID NO:85) derived from a hidden Markov model.

[0433] To identify the presence of a "PAS" domain in a 33751 protein sequence, and make the determination that a

polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against a SMART database (Simple Modular Architecture Research Tool) of HMMs as described in Schultz et al. (1998), Proc. Natl. Acad. Sci. USA 95:5857 and Schultz et al. (200) Nucl. Acids Res 28:231. The database contains domains identified by profiling with the hidden Markov models of the HMMer2 search program (R. Durbin et al. (1998) Biological sequence analysis: probabilistic models of proteins and nucleic acids. Cambridge University Press). The database also is extensively annotated and monitored by experts to enhance accuracy. A search was performed against the HMM database resulting in the identification of a "PAS" domain in the amino acid sequence of human 33751 at about residues 41 to 60 of SEQ ID NO:82.

[0434] In a preferred embodiment, a 33751 polypeptide or protein has a "PAS domain" or a region which includes at least about 10 to 100, more preferably about 13 to 50, or 17 to 25 amino acid residues and has at least about 50%, 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "PAS domain," e.g., the PAS domain of human 33751 (e.g., residues 41 to 60 of SEQ ID NO:82).

[0435] A 33751 polypeptide can further include a "PAC domain" or regions homologous with a "PAC domain." As used herein, a "PAC domain" includes an amino acid sequence of about 10 to 100 amino acid residues in length. Preferably, the PAC domain contributes to the folding of the PAS domain. Preferably, the PAC domain is located at the C-terminal end of the PAS domain in a 33751 polypeptide. The term "PAC domain" includes an amino acid sequence of about 10 to 100 amino acid residues in length and having a bit score for the alignment of the sequence to the cyclic nucleotide binding domain (HMM) of at least 10. Preferably, a PAC domain includes at least about 10 to 100 amino acids, more preferably about 20 to 50 amino acid residues, or about 25 to 30 amino acids and has a bit score for the alignment of the sequence to the PAC domain (HMM) of at least 15 or greater. The PAC domain (HMM) has been assigned the PFAM Accession PF00785. The PAC domain (amino acids 93 to 120 of SEQ ID NO:82) of human 33751 aligns with a consensus amino acid sequence (SEQ ID NO:86) derived from a hidden Markov model.

[0436] To identify the presence of a "PAC" domain in a 33751 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against a SMART database (Simple Modular Architecture Research Tool) of HMMs as described in Schultz et al. (1998), Proc. Natl. Acad. Sci. USA 95:5857 and Schultz et al. (200) Nucl. Acids Res 28:231. The database contains domains identified by profiling with the hidden Markov models of the HMMer2 search program (R. Durbin et al. (1998) Biological sequence analysis: probabilistic models of proteins and nucleic acids. Cambridge University Press). The database also is extensively annotated and monitored by experts to enhance accuracy. A search was performed against the HMM database resulting in the identification of a "PAC" domain in the amino acid sequence of human 33751 at about residues 93 to 120 of SEQ ID NO:82.

[0437] In a preferred embodiment, a 33751 polypeptide or protein has a "PAC domain" or a region which includes at least about 10 to 100, 20 to 50, or 25 to 30 amino acid residues and has at least about 50%, 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "PAC domain," e.g., the PAC domain of human 33751 (e.g., residues 93 to 120 of SEQ ID NO:82).

[0438] A 33751 molecule can further include a cyclic nucleotide binding domain or regions homologous with a "cyclic nucleotide binding domain." As used herein, the term "cyclic nucleotide binding domain" includes an amino acid sequence of about 50 to 200 amino acid residues in length and having a bit score for the alignment of the sequence to the cyclic nucleotide binding domain (HMM) of at least 50. Preferably, a cyclic nucleotide binding domain is capable of binding a cyclic nucleotide (e.g., cAMP or cGMP), and is composed of three 1-helices and a distinctive eight-stranded antiparallel J-barrel structure. Preferably, a cyclic nucleotide binding domain includes at least about 50 to 200 amino acids, more preferably about 70 to 120 amino acid residues, or about 85 to 95 amino acids and has a bit score for the alignment of the sequence to the cyclic nucleotide binding domain (HMM) of at least 75 or greater. The cyclic nucleotide binding domain (HMM) has been assigned the PFAM Accession PF00027. The cyclic nucleotide binding domain (amino acids 760 to 850 of SEQ ID NO:82) of human 33751 aligns with a consensus amino acid sequence (SEQ ID NO:87) derived from a hidden Markov model.

[0439] To identify the presence of a "cyclic nucleotide binding" domain in a 33751 protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against a SMART database (Simple Modular Architecture Research Tool) of HMMs as described in Schultz et al. (1998), *Proc. Natl. Acad. Sci. USA* 95:5857 and Schultz et al. (200) Nucl. Acids Res 28:231. The database contains domains identified by profiling with the hidden Markov models of the HMMer2 search program (R. Durbin et al. (1998) *Biological sequence analysis: probabilistic models of proteins and nucleic acids*. Cambridge University Press.) The database also is extensively annotated and monitored by experts to enhance accuracy. A search was performed against the HMM database resulting in the identification of a "cyclic nucleotide binding" domain in the amino acid sequence of human 33751 at about residues 760 to 850 of SEQ ID NO:82.

[0440] In a preferred embodiment a 33751 polypeptide or protein has a "cyclic nucleotide binding domain" or a region which includes at least about 50 to 200, more preferably about 70 to 120, or 85 to 95 amino acid residues and has at least about 50%, 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "cyclic nucleotide binding domain," e.g., the cyclic nucleotide binding domain of human 33751 (e.g., residues 760 to 850 of SEQ ID NO:82).

[0441] A 33751 protein further includes a predicted N-terminal cytoplasmic domain located at about amino acids 1-411 of SEQ ID NO:82. As used herein, a "N-terminal cytoplasmic domain" includes an amino acid sequence having about 1 to 600, preferably about 1 to 500, or even more preferably about 1 to 420 amino acid residues in length and is located inside of a cell or intracellularly. The C-terminal amino acid residue of a "N-terminal cytoplasmic domain" is adjacent to a N-terminal amino acid residue of a transmembrane domain in a 33751 protein. For example, a N-terminal cytoplasmic domain is located at about amino acid residues 1 to 411 of SEQ ID NO:82.

[0442] In a preferred embodiment 33751 polypeptide or protein has an "N-terminal cytoplasmic domain" or a region which includes at least about 1 to 600, preferably about 100 to

420, and even more preferably about 411 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with an "N-terminal cytoplasmic domain," e.g., the N-terminal cytoplasmic domain of human 33751 (e.g., residues (1 to 411 of SEQ ID NO:82).

[0443] In another embodiment, a 33751 protein includes a "C-terminal cytoplasmic domain," also referred to herein as a C-terminal cytoplasmic tail, in the sequence of the protein. As used herein, a "C-terminal cytoplasmic domain" includes an amino acid sequence having a length of at least about 200, more preferably 400 or more amino acid residues and is located within a cell or within the cytoplasm of a cell. Accordingly, the N-terminal amino acid residue of a "C-terminal cytoplasmic domain" is adjacent to a C-terminal amino acid residue of a transmembrane domain in a naturally-occurring 33751 protein. For example, a C-terminal cytoplasmic domain is found at about amino acid residues 667 to 1196 of SEQ ID NO:82.

[0444] In a preferred embodiment, a 33751 polypeptide or protein has a C-terminal cytoplasmic domain or a region which includes at least about 200, more preferably 400 or more amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with an "C-terminal cytoplasmic domain," e.g., the C-terminal cytoplasmic domain of human 33751 (e.g., residues 667 to 1196 of SEQ ID NO:82).

[0445] 33751 proteins can further include at least one, two, three, four, five, and preferably six transmembrane domains. As used herein, the term "transmembrane domain" includes an amino acid sequence of about 10 to 45, preferably 12 to 30, and most preferably 15 to 25, amino acid residues in length that spans the plasma membrane. More preferably, a transmembrane domain includes about at least 17, 18, 19, 22, or 25 amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an alpha-helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, e.g., leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains are described in, for example, Zagotta W. N. et al, (1996) Annual Rev. Neurosci. 19: 235-263, the contents of which are incorporated herein by reference. Amino acid residues 412 to 433, 453 to 470, 495 to 513, 549 to 573, 614 to 630, and 642 to 666 of SEQ ID NO:82 are transmembrane domains. Accordingly, proteins having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, about 80-90%, or about 90-100% homology with amino acids 412 to 433, 453 to 470, 495 to 513, 549 to 573, 614 to 630, and 642 to 666 of SEQ ID NO:82 are within the scope of the invention.

[0446] In another embodiment, a 33751 protein includes at least one, or two cytoplasmic loop, also referred to herein as a cytoplasmic domain. As used herein, a "cytoplasmic loop" includes an amino acid sequence having a length of at least about 10, preferably about 20, amino acid residues located within a cell or within the cytoplasm of a cell. For example, a cytoplasmic loop is found at about amino acids 471 to 494, or 574 to 613 of SEQ ID NO:82.

[0447] In a preferred embodiment 33751 polypeptide or protein has at least one cytoplasmic loop or a region which includes at least about 10, preferably about 20 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with an "cytoplasmic loop," e.g., at least one cytoplasmic loop of human 33751 (e.g., residues 471 to 494, or 574 to 613 of SEQ ID NO:82).

[0448] In another embodiment, a 33751 protein include at least one, two, or three extracellular loop. As defined herein, the term "loop" includes an amino acid sequence that resides outside of a phospholipid membrane, having a length of at least about 20 to 70, and preferably about 30 to 50 amino acid residues, and has an amino acid sequence that connects two transmembrane domains within a protein or polypeptide. Extracellular domains are located outside of the cell. Accordingly, the N-terminal amino acid of a non-cytoplasmic loop is adjacent to a C-terminal amino acid of a transmembrane domain in a 33751 protein, and the C-terminal amino acid of a non-cytoplasmic loop is adjacent to an N-terminal amino acid of a transmembrane domain in a 33751 protein. For example, an "extracellular loop" can be found at about amino acids 434 to 452, 514 to 548, and 631 to 641 of SEQ ID NO:82.

[0449] In a preferred embodiment, a 33751 polypeptide or protein has at least one, two, or three extracellular loops or regions which include at least about 5, preferably about 5 to 80, and more preferably about 20 to 50 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with an "non-cytoplasmic loop," e.g., at least one non-cytoplasmic loop of human 33751 (e.g., residues 403 to 433, 493 to 540, and 604 to 645 of SEQ ID NO:82).

[0450] Accordingly, in one embodiment of the invention, a 33751 includes at least one, two, three, four, five, preferably six, transmembrane domains, at least one, or two cytoplasmic loops, and/or at least one, two, or three non-cytoplasmic loops. In another embodiment, the 33751 further includes an N-terminal and a C-terminal cytoplasmic domains.

[0451] A 33751 family member can include at least one predicted ion transport protein domain, at least one predicted PAS domain, at least one predicted PAC domain, and at least one predicted cyclic nucleotide-binding domain. Furthermore, a 33751 family member can include at least one, two, three, four, five, six, seven, eight, nine, or preferably ten predicted N-glycosylation sites (PS00001); at least one, two, three, or preferably four predicted cAMP- and cGMP-dependent protein kinase phosphorylation sites (PS00004); at least one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, or preferably twenty-one predicted protein kinase C phosphorylation sites (PS00005); at least one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, twenty-one, twenty-two, twenty-three, or preferably twenty-four predicted casein kinase II phosphorylation sites (PS00006); at least one, two, or preferably three tyrosine kinase phosphorylation sites (PS00007); at least one, two, three, four, five, six, seven, eight, nine, or preferably ten predicted N-myristylation sites (PS00008); and at least one predicted amidation site (PS00009).

[0452] As the 33751 polypeptides of the invention may modulate 33751-mediated activities, they may be useful as of for developing novel diagnostic and therapeutic agents for 33751-mediated or related disorders, as described below.

[0453] As used herein, a "33751 activity," "biological activity of 33751," or "functional activity of 33751," refers to an activity exerted by a 33751 protein, polypeptide or nucleic acid molecule on e.g., a 33751-responsive cell or on a 33751 substrate, e.g., a protein substrate, as determined in vivo or in

vitro. In one embodiment, a 33751 activity is a direct activity, such as an association with a 33751 target molecule. A "target molecule" "substrate" or "binding partner" is a molecule with which a 33751 protein binds or interacts in nature. A 33751 activity can also be an indirect activity, e.g., a cellular signaling activity mediated by interaction of the 33751 protein with a 33751-binding partner. In an exemplary embodiment, 33751 is controlling one or more of membrane excitability, and/or the frequency and pattern of neuronal firing.

[0454] Based on the above-described sequence similarities and the tissue distribution described below, the 33751 molecules of the present invention are predicted to have similar biological activities as potassium channel family members. Thus, a 33751 potassium channel or subsequence or variant polypeptide may have one or more of the aforesaid domains and, therefore, one or more activities or functions characteristic of a potassium channel family member, including, but not limited to, (1) controlling neurotransmitter release from neurons; (2) modulating repolarization of the neuronal cell membrane; (3) contributing to the formation of voltage-gated potassium channels; (4) binding to cyclic nucleotides; (5) regulating nociceptive responses; (6) regulating synaptic transmission; (7) modulating pain or inflammation response; or (8) regulating the frequency and pattern of neuronal firing. Thus, the 33751 molecules can act as novel diagnostic targets and therapeutic agents for controlling potassium channel associated disorders.

[0455] Activation of K+ channels affects the frequency and the pattern of neuronal, firing. Several voltage-gated K+ channels are expressed in subpopulation of sensory neurons including those involved in nociception. It has been shown that the expression of some voltage-gated K+ channels decreases in dorsal root ganglion neurons after axotomy, and that the peak of K+ currents is reduced in sensory neurons during chronic inflammation. Furthermore, administration of K+ channel openers potentiated the antinociception produced by agonists of 1-2-adrenoreceptors or by morphin.

[0456] TaqMan experiments show high levels of 33751 mRNA expression in the human brain, followed by the dorsal root ganglion (DRG) and spinal cord (Table 16). A TaqMan experiment detecting mRNA expression of the rat ortholog of human 33751 revealed a similar pattern of expression to that of the human gene. The results indicate that 33751 gene is a nervous system specific gene. Further, TaqMan experiments in a rat model show down-regulation of 33751 mRNA in the DRG after CCI and after axotomy. TaqMan experiments in the rat model show down-regulation of 33751 mRNA in spinal cord after CFA injection and after axotomy (Tables 17-19). In situ hybridization with a human probe shows the expression of 33751 mRNA in monkey and rat brain, spinal cord, and DRG. In the monkey cord, 33751 mRNA is expressed in lamina V in large size neurons, most likely spinothalamic neurons. In the DRG, a small subpopulation of neurons expressed high levels of 33751 mRNA. Another subpopulation of sensory neurons expressed much lower levels of 33751 mRNA. Down-regulation of 33751 was observed by in situ hybridization 14 and 28 days after axotomy. Accordingly, 33751 may be critical for hypersensitivity in different pain states, and thus may represent a unique target for pain.

[0457] Animal models of pain response include, but are not limited to, axotomy, the cutting or severing of an axon; chronic constriction injury (CCI), a model of neuropathic pain which involves ligation of the sciatic nerve in rodents, e.g., rats; or intraplantar Freund's adjuvant injection as a model of arthritic pain. Other animal models of pain response are described in, e.g., *ILAR Journal* (1999) Volume 40, Number 3 (entire issue). TaqMan experiments in rat animal models show no regulation in DRGs. However, 33751 mRNA is up-regulated in the spinal cord after CCI axotomy, and after CFA intraplantar injection. These experiments indicate a role for the 33751 molecule in pain response.

[0458] Therefore, 33751-associated disorders can detrimentally affect regulation and modulation of the pain response; and vasoconstriction response and pain therefrom. Examples of 33751 associated disorders in which the 33751 molecules of the invention may be directly or indirectly involved include pain, pain syndromes, and inflammatory disorders, including inflammatory pain, and therefore, modulators of the activity or expression of 33751 polypeptides may be useful for developing novel diagnostic and therapeutic agents for controlling pain, pain disorders, and inflammatory disorders.

[0459] Agents that modulate 33751 polypeptide or nucleic acid activity or expression can be used to treat pain elicited by any medical condition. A subject receiving the treatment can be additionally treated with a second agent, e.g., an anti-inflammatory agent, an antibiotic, or a chemotherapeutic agent, to further ameliorate the condition.

[0460] The 33751 molecules can also act as novel diagnostic targets and therapeutic agents controlling pain caused by other disorders, e.g., cancer, e.g., prostate cancer. Accordingly, the 33751 molecules can act as novel diagnostic targets and therapeutic agents for controlling one or more of cellular proliferative and/or differentiative disorders, or pain therefrom.

[0461] As the 33751 mRNA is highly expressed in the brain, 33751 molecules can also act as novel diagnostic targets and therapeutic agents for brain disorders.

Tissue Distribution of 33751 mRNA by TaqMan Analysis

[0462] A TaqMan experiment shows 33751 mRNA was highly expressed in brain, followed by DRG and spinal cord. The relative tissue distribution of 33751 mRNA is depicted in tabular form in Table 16.

[0463] The relative tissue distribution of 33751 mRNA using rat panels is depicted in tabular form in Tables 17-19. Expression of rat 33751 mRNA was observed in DRG, brain, SCG, spinal, optic nerve, and thyroid. A TaqMan experiment in an additional rat panel shows down-regulation of 33751 in DRG after CCI and axotomy. A final TaqMan experiment in a rat panel shows down-regulation of 33751 in spinal cord after CFA injection and after axotomy. The relative tissue distribution of 33751 mRNA is depicted in tabular form in Tables 17-19.

TABLE 16

Tissue Type	33751	β2.803	∂Ct	Expression
Adrenal Gland	38.89	18.04	20.86	0.00
Brain	25.27	20.21	5.06	29.98
Heart	40.00	18.22	21.78	0.00
Kidney	37.24	18.34	18.90	0.00
Liver	39.98	18.78	21.20	0.00
Lung	40.00	16.48	23.52	0.00
Mammary Gland	40.00	17.60	22.41	0.00
Pancreas	38.25	20.64	17.61	0.00
Placenta	37.57	18.09	19.48	0.00
Prostate	40.00	17.47	22.54	0.00
Salivary Gland	40.00	18.52	21.49	0.00
Muscle	40.00	20.29	19.71	0.00

TABLE 16-continued

Tissue Type	33751	β2.803	∂Ct	Expression
Sm. Intestine	38.10	18.48	19.62	0.00
Spleen	40.00	16.29	23.72	0.00
Stomach	32.67	17.76	14.91	0.03
Teste	34.00	18.88	15.13	0.03
Thymus	35.32	17.23	18.09	0.00
Trachea	40.00	18.32	21.68	0.00
Uterus	40.00	18.56	21.45	0.00
Spinal Cord	30.83	18.84	11.99	0.25
DRG	29.89	19.06	10.83	0.55
Skin	40.00	18.18	21.83	0.00

TABLE 17

Tissue	r33751	18S	∂Ct	Expression
Brain	35.23	12.59	22.64	0.01
Spinal Cord	36.91	12.39	24.52	0.00
DRG	33.27	12.07	21.20	0.03
SCG	35.79	13.00	22.79	0.01
Optic Nerve	37.44	13.23	24.21	0.00
Hairy Skin	40.00	14.06	25.94	0.00
Gastro Muscle	39.68	12.39	27.29	0.00
Heart	40.00	13.29	26.71	0.00
Kidney	40.00	13.12	26.88	0.00
Liver	40.00	12.61	27.39	0.00
Lung	40.00	13.48	26.53	0.00
Spleen	40.00	13.12	26.88	0.00
Aorta	40.00	13.33	26.67	0.00
Adrenal Gland	40.00	12.75	27.25	0.00
Salivary Gland	40.00	12.51	27.50	0.00
Thyroid	35.88	12.46	23.43	0.01
Prostate	40.00	13.52	26.48	0.00
Thymus	40.00	12.69	27.32	0.00
Trachea	40.00	13.92	26.09	0.00
Esophagus	40.00	12.86	27.14	0.00
Duodenum	39.97	13.92	26.06	0.00
Diaphragm	40.00	14.34	25.66	0.00
Colon	40.00	15.05	24.96	0.00

TABLE 18

Tissue	r33751	18S	∂Ct	Expression
Naïve DRG	33.84	12.83	21.01	0.005
I DRG CCI 3	33.61	13.00	20.61	0.006
I DRG CCI 7	35.42	13.08	22.34	0.002
I DRG CCI 10	36.74	12.98	23.77	0.001
I DRG CCI 14	36.95	12.77	24.18	0.001
I DRG CCI 28	36.64	12.76	23.88	0.001
Naïve DRG	33.88	12.86	21.02	0.005
I DRG CFA 1	33.23	12.92	20.31	0.008
I DRG CFA 3	34.59	13.01	21.58	0.003
I DRG CFA 7	34.19	12.57	21.62	0.003
I DRG CFA 10	33.99	12.61	21.38	0.004
I DRG CFA 14	34.05	12.79	21.27	0.004
I DRG CFA 28	33.34	13.13	20.21	0.008
Naïve DRG	33.80	12.73	21.07	0.005
I DRG AXT 1	32.98	12.45	20.53	0.007
I DRG AXT 3	34.30	13.35	20.96	0.005
I DRG AXT 7	36.17	12.74	23.43	0.001
I DRG AXT 14	35.32	12.54	22.78	0.001

TABLE 19

Tissue	r33751	18S	∂Ct	Expression
Naïve SC	36.49	13.55	22.94	0.001
I SC CCI 3	37.02	13.56	23.46	0.001

TABLE 19-continued

Tissue	r33751	18S	∂Ct	Expression
I SC CCI 7	36.66	13.56	23.11	0.001
I SC CCI 10	36.14	13.76	22.38	0.002
I SC CCI 14	36.96	13.26	23.70	0.001
I SC CCI 28	38.19	13.34	24.85	0.000
Naïve SC	38.79	13.10	23.98	0.001
I SC CFA 1	37.10	13.73	23.38	0.001
I SC CFA 3	36.24	13.18	23.07	0.001
I SC CFA 7	36.66	12.73	23.93	0.001
I SC CFA 10	37.94	13.26	24.68	0.000
I SC CFA 14	39.17	14.04	25.13	0.000
I SC CFA 28	37.23	13.79	23.45	0.001
Naïve SC	36.76	13.61	23.15	0.001
I SC AXT 1	36.13	13.50	22.63	0.002
I SC AXT 3	38.53	13.16	25.37	0.000
I SC AXT 7	38.02	13.11	24.92	0.000
I SC AXT 14	38.84	14.04	24.80	0.000

Electrophysiological Characterization of 33751

[0464] Electrophysiological characterization of the 33751 polypeptide was performed. The characteristic pattern of voltage-dependent inactivation of the human 33751 and its rat ortholog was found (Shi, W. et al. (1997) *J. Neurosci*. Vol. 17(24):9423). The conductance of 33751-expressed in CHOK1 cells maintained moderate depolarizations below −10 mV. 33751 polypeptide showed sustained currents without inactivation. The effect of dofetillide on the membrane potential of 911 cells transiently transfected with 33751 versus vector alone was also studied. In this experiment, 911 cells were placed at 8,000 cells per well, and loaded with a MP dye for 40 minutes at 37° C. 10×30 mM KCl or 0.3-3 uM dofetilide (made up in assay buffer) was added after 3 minute when baseline was read.

DEFINITIONS

[0465] The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, fragments thereof, and derivatives and other variants of the sequence in SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 thereof are collectively referred to as "polypeptides or proteins of the invention" or "18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptides or proteins". Nucleic acid molecules encoding such polypeptides or proteins are collectively referred to as "nucleic acids of the invention" or "18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acids."

[0466] As used herein, the term "nucleic acid molecule" includes DNA molecules (e.g., a cDNA or genomic DNA) and RNA molecules (e.g., an mRNA) and analogs of the DNA or RNA generated, e.g., by the use of nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

[0467] The term "isolated or purified nucleic acid molecule" includes nucleic acid molecules which are separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. For example, with regards to genomic DNA, the term "isolated" includes nucleic acid molecules which are separated from the chromosome with which the genomic DNA is naturally associated. Preferably,

an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and/or 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of 5' and/or 3' nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

[0468] As used herein, the term "hybridizes under low stringency, medium stringency, high stringency, or very high stringency conditions" describes conditions for hybridization and washing. Guidance for performing hybridization reactions can be found in *Current Protocols in Molecular Biology* (1989) John Wiley & Sons, N.Y., 6.3.1-6.3.6, which is incorporated by reference. Aqueous and nonaqueous methods are described in that reference and either can be used. Specific hybridization conditions referred to herein are as follows: 1) low stringency hybridization conditions in 6× sodium chloride/sodium citrate (SSC) at about 45° C., followed by two washes in 0.2×SSC, 0.1% SDS at least at 50° C. (the temperature of the washes can be increased to 55° C. for low stringency conditions); 2) medium stringency hybridization conditions in 6×SSC at about 45° C., followed by one or more washes in 0.2×SSC, 0.1% SDS at 60° C.; 3) high stringency hybridization conditions in 6×SSC at about 45° C., followed by one or more washes in 0.2×SSC, 0.1% SDS at 65° C.; and preferably 4) very high stringency hybridization conditions are 0.5M sodium phosphate, 7% SDS at 65° C., followed by one or more washes at 0.2×SSC, 1% SDS at 65° C. Very high stringency conditions (4) are the preferred conditions and the ones that should be used unless otherwise specified.

[0469] As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

[0470] As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules which include an open reading frame encoding a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, preferably a mammalian 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, and can further include non-coding regulatory sequences, and introns.

[0471] An "isolated" or "purified" polypeptide or protein is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. In one embodiment, the language "substantially free" means preparation of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein having less than about 30%, 20%, 10% and more preferably 5% (by dry weight), of non-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein (also referred to herein as a "contaminating protein"), or of chemical precursors or non-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 chemicals. When the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation. The invention includes isolated or purified preparations of at least 0.01, 0.1, 1.0, and 10 milligrams in dry weight.

[0472] A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 (e.g., the sequence of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83) without abolishing or more preferably, without substantially altering a biological activity, whereas an "essential" amino acid residue results in such a change. For example, amino acid residues that are conserved among the polypeptides of the present invention, e.g., those present in the conserved domains, are predicted to be particularly unamenable to alteration.

[0473] 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., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

[0474] As used herein, a "biologically active portion" of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein includes a fragment of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein which participates in an interaction between a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecule and a non-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914,

59921 or 33751 molecule. Biologically active portions of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, e.g., the amino acid sequence shown in SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, which include fewer amino acids than the full length 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, and exhibit at least one activity of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. A biologically active portion of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein can be a polypeptide which is, for example, 10, 25, 50, 100, 200 or more amino acids in length. Biologically active portions of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein can be used as targets for developing agents which modulate a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mediated activity.

[0475] Calculations of homology or sequence identity (the terms "homology" and "identity" are used interchangeably herein) between sequences are performed as follows:

[0476] 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-homologous 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%, 90%, 100% of the length of the reference sequence. 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.

[0477] 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 (1970) *J. Mol. Biol.* 48:444-453 algorithm which has been incorporated into the GAP program in the GCG software package using either a Blossum 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. A particularly preferred set of parameters (and the one that should be used if the practitioner is uncertain about what parameters should be applied to determine if a molecule is within a sequence identity or homology limitation of the invention) are a Blossum 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

[0478] The percent identity between two amino acid or nucleotide sequences can be determined using the algorithm of Meyers and Miller ((1989) CABIOS, 4:11-17) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

[0479] The nucleic acid and protein sequences described herein can be used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul et al. (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score=100, wordlength=12 to obtain nucleotide sequences homologous to 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score=50, wordlength=3 to obtain amino acid sequences homologous to 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) *Nucleic Acids Res.* 25:3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used.

[0480] Particular 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptides of the present invention have an amino acid sequence substantially identical to the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82. In the context of an amino acid sequence, the term "substantially identical" is used herein to refer to a first amino acid that contains a sufficient or minimum number of amino acid residues that are i) identical to, or ii) conservative substitutions of aligned amino acid residues in a second amino acid sequence such that the first and second amino acid sequences can have a common structural domain and/or common functional activity. For example, amino acid sequences that contain a common structural domain having at least about 60%, or 65% identity, likely 75% identity, more likely 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 are termed substantially identical.

[0481] In the context of nucleotide sequence, the term "substantially identical" is used herein to refer to a first nucleic acid sequence that contains a sufficient or minimum number of nucleotides that are identical to aligned nucleotides in a second nucleic acid sequence such that the first and second nucleotide sequences encode a polypeptide having common

functional activity, or encode a common structural polypeptide domain or a common functional polypeptide activity. For example, nucleotide sequences having at least about 60%, or 65% identity, likely 75% identity, more likely 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83 are termed substantially identical.

[0482] "Misexpression or aberrant expression", as used herein, refers to a non-wild type pattern of gene expression, at the RNA or protein level. It includes: expression at non-wild type levels, i.e., over or under expression; a pattern of expression that differs from wild type in terms of the time or stage at which the gene is expressed, e.g., increased or decreased expression (as compared with wild type) at a predetermined developmental period or stage; a pattern of expression that differs from wild type in terms of decreased expression (as compared with wild type) in a predetermined cell type or tissue type; a pattern of expression that differs from wild type in terms of the splicing size, amino acid sequence, post-transitional modification, or biological activity of the expressed polypeptide; a pattern of expression that differs from wild type in terms of the effect of an environmental stimulus or extracellular stimulus on expression of the gene, e.g., a pattern of increased or decreased expression (as compared with wild type) in the presence of an increase or decrease in the strength of the stimulus.

[0483] "Subject", as used herein, can refer to a mammal, e.g., a human, or to an experimental or animal or disease model. The subject can also be a non-human animal, e.g., a horse, cow, goat, or other domestic animal.

[0484] A "purified preparation of cells", as used herein, refers to, in the case of plant or animal cells, an in vitro preparation of cells and not an entire intact plant or animal. In the case of cultured cells or microbial cells, it consists of a preparation of at least 10% and more preferably 50% of the subject cells.

[0485] As used herein, cellular proliferative and/or differentiative disorders include cancer, e.g., carcinoma, sarcoma, metastatic disorders or hematopoietic neoplastic disorders, e.g., leukemias. A metastatic tumor can arise from a multitude of primary tumor types, including but not limited to those of prostate, colon, lung, breast and liver origin.

[0486] As used herein, a "cellular proliferation, growth, differentiation, or migration process" is a process by which a cell increases in number, size or content, by which a cell develops a specialized set of characteristics which differ from that of other cells, or by which a cell moves closer to or further from a particular location or stimulus.

[0487] As used herein, the term "cancer" (also used interchangeably with the terms, "hyperproliferative" and "neoplastic") refers to cells having the capacity for autonomous growth, i.e., an abnormal state or condition characterized by rapidly proliferating cell growth. Cancerous disease states may be categorized as pathologic, i.e., characterizing or constituting a disease state, e.g., malignant tumor growth, or may be categorized as non-pathologic, i.e., a deviation from normal but not associated with a disease state, e.g., cell proliferation associated with wound repair. The term is meant to include all types of cancerous growths or oncogenic processes, metastatic tissues or malignantly transformed cells, tissues, or organs, irrespective of histopathologic type or stage of invasiveness. The term "cancer" includes malignancies of the various organ systems, such as those affecting lung, breast, thyroid, lymphoid, gastrointestinal, and genitourinary tract, as well as adenocarcinomas which include malignancies such as most colon cancers, renal-cell carcinoma, prostate cancer and/or testicular tumors, non-small cell carcinoma of the lung, cancer of the small intestine and cancer of the esophagus. The term "carcinoma" is art recognized and refers to malignancies of epithelial or endocrine tissues including respiratory system carcinomas, gastrointestinal system carcinomas, genitourinary system carcinomas, testicular carcinomas, breast carcinomas, prostatic carcinomas, endocrine system carcinomas, and melanomas. Exemplary carcinomas include those forming from tissue of the cervix, lung, prostate, breast, head and neck, colon and ovary. The term "carcinoma" also includes carcinosarcomas, e.g., which include malignant tumors composed of carcinomatous and sarcomatous tissues. An "adenocarcinoma" refers to a carcinoma derived from glandular tissue or in which the tumor cells form recognizable glandular structures. The term "sarcoma" is art recognized and refers to malignant tumors of mesenchymal derivation.

[0488] Examples of cellular proliferative and/or differentiative disorders of the lung include, but are not limited to, tumors such as bronchogenic carcinoma, including paraneoplastic syndromes, bronchioloalveolar carcinoma, neuroendocrine tumors, such as bronchial carcinoid, miscellaneous tumors, metastatic tumors, and pleural tumors, including solitary fibrous tumors (pleural fibroma) and malignant mesothelioma.

[0489] Examples of cellular proliferative and/or differentiative disorders of the breast include, but are not limited to, proliferative breast disease including, e.g., epithelial hyperplasia, sclerosing adenosis, and small duct papillomas; tumors, e.g., stromal tumors such as fibroadenoma, phyllodes tumor, and sarcomas, and epithelial tumors such as large duct papilloma; carcinoma of the breast including in situ (noninvasive) carcinoma that includes ductal carcinoma in situ (including Paget's disease) and lobular carcinoma in situ, and invasive (infiltrating) carcinoma including, but not limited to, invasive ductal carcinoma, invasive lobular carcinoma, medullary carcinoma, colloid (mucinous) carcinoma, tubular carcinoma, and invasive papillary carcinoma, and miscellaneous malignant neoplasms. Disorders in the male breast include, but are not limited to, gynecomastia and carcinoma.

[0490] Examples of cellular proliferative and/or differentiative disorders involving the colon include, but are not limited to, tumors of the colon, such as non-neoplastic polyps, adenomas, familial syndromes, colorectal carcinogenesis, colorectal carcinoma, and carcinoid tumors.

[0491] Examples of cancers or neoplastic conditions, in addition to the ones described above, include, but are not limited to, a fibrosarcoma, myosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma, angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, gastric cancer, esophageal cancer, rectal cancer, pancreatic cancer, ovarian cancer, prostate cancer, uterine cancer, cancer of the head and neck, skin cancer, brain cancer, squamous cell carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinoma, cystadenocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hepatoma, bile duct carcinoma, choriocarcinoma, seminoma, embryonal carcinoma, Wilm's tumor, cervical cancer, testicular cancer, small cell lung carcinoma, non-small cell lung

carcinoma, bladder carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodendroglioma, meningioma, melanoma, neuroblastoma, retinoblastoma, leukemia, lymphoma, or Kaposi sarcoma.

[0492] Proliferative disorders include hematopoietic neoplastic disorders. As used herein, the term "hematopoietic neoplastic disorders" includes diseases involving hyperplastic/neoplastic cells of hematopoietic origin, e.g., arising from myeloid, lymphoid or erythroid lineages, or precursor cells thereof. Preferably, the diseases arise from poorly differentiated acute leukemias, e.g., erythroblastic leukemia and acute megakaryoblastic leukemia. Additional exemplary myeloid disorders include, but are not limited to, acute promyeloid leukemia (APML), acute myelogenous leukemia (AML) and chronic myelogenous leukemia (CML) (reviewed in Vaickus (1991) *Crit. Rev. in Oncol./Hemotol.* 11:267-97); lymphoid malignancies include, but are not limited to acute lymphoblastic leukemia (ALL) which includes B-lineage ALL and T-lineage ALL, chronic lymphocytic leukemia (CLL), prolymphocytic leukemia (PLL), hairy cell leukemia (HLL) and Waldenstrom's macroglobulinemia (WM). Additional forms of malignant lymphomas include, but are not limited to non-Hodgkin lymphoma and variants thereof, peripheral T cell lymphomas, adult T cell leukemia/lymphoma (ATL), cutaneous T-cell lymphoma (CTCL), large granular lymphocytic leukemia (LGF), Hodgkin's disease and Reed-Sternberg disease.

[0493] As used herein, an "angiogenic or angiogenesis disorder" includes a disease or disorder which affects or is caused by aberrant or deficient angiogenesis. Disorders involving angiogenesis include, but are not limited to, aberrant or excess angiogenesis in tumors such as hemangiomas and Kaposi's sarcoma, von Hippel-Lindau disease, as well as the angiogenesis associated with tumor growth; aberrant or excess angiogenesis in diseases such as a Castleman's disease or fibrodysplasia ossificans progressiva; aberrant or deficient angiogenesis associated with aging, complications of healing certain wounds and complications of diseases such as diabetes and rheumatoid arthritis; or aberrant or deficient angiogenesis associated with hereditary hemorrhagic telangiectasia, autosomal dominant polycystic kidney disease, myelodysplastic syndrome or Klippel-Trenaunay-Weber syndrome.

[0494] As used herein, disorders of the breast include, but are not limited to, disorders of development; inflammations, including but not limited to, acute mastitis, periductal mastitis, periductal mastitis (recurrent subareolar abscess, squamous metaplasia of lactiferous ducts), mammary duct ectasia, fat necrosis, granulomatous mastitis, and pathologies associated with silicone breast implants; fibrocystic changes; proliferative breast disease including, but not limited to, epithelial hyperplasia, sclerosing adenosis, and small duct papillomas; tumors including, but not limited to, stromal tumors such as fibroadenoma, phyllodes tumor, and sarcomas, and epithelial tumors such as large duct papilloma; carcinoma of the breast including in situ (noninvasive) carcinoma that includes ductal carcinoma in situ (including Paget's disease) and lobular carcinoma in situ, and invasive (infiltrating) carcinoma including, but not limited to, invasive ductal carcinoma, no special type, invasive lobular carcinoma, medullary carcinoma, colloid (mucinous) carcinoma, tubular carcinoma, and invasive papillary carcinoma, and miscellaneous malignant neoplasms. Disorders in the male breast include, but are not limited to, gynecomastia and carcinoma.

[0495] As used herein, disorders involving the colon include, but are not limited to, congenital anomalies, such as atresia and stenosis, Meckel diverticulum, congenital aganglionic megacolon-Hirschsprung disease; enterocolitis, such as diarrhea and dysentery, infectious enterocolitis, including viral gastroenteritis, bacterial enterocolitis, necrotizing enterocolitis, antibiotic-associated colitis (pseudomembranous colitis), and collagenous and lymphocytic colitis, miscellaneous intestinal inflammatory disorders, including parasites and protozoa, acquired immunodeficiency syndrome, transplantation, drug-induced intestinal injury, radiation enterocolitis, neutropenic colitis (typhlitis), and diversion colitis; idiopathic inflammatory bowel disease, such as Crohn disease and ulcerative colitis; tumors of the colon, such as non-neoplastic polyps, adenomas, familial syndromes, colorectal carcinogenesis, colorectal carcinoma, and carcinoid tumors.

[0496] As used herein, disorders involving the kidney, also referred to herein as renal disorders, include, but are not limited to, congenital anomalies including, but not limited to, cystic diseases of the kidney, that include but are not limited to, cystic renal dysplasia, autosomal dominant (adult) polycystic kidney disease, autosomal recessive (childhood) polycystic kidney disease, and cystic diseases of renal medulla, which include, but are not limited to, medullary sponge kidney, and nephronophthisis-uremic medullary cystic disease complex, acquired (dialysis-associated) cystic disease, such as simple cysts; glomerular diseases including pathologies of glomerular injury that include, but are not limited to, in situ immune complex deposition, that includes, but is not limited to, anti-GBM nephritis, Heymann nephritis, and antibodies against planted antigens, circulating immune complex nephritis, antibodies to glomerular cells, cell-mediated immunity in glomerulonephritis, activation of alternative complement pathway, epithelial cell injury, and pathologies involving mediators of glomerular injury including cellular and soluble mediators, acute glomerulonephritis, such as acute proliferative (poststreptococcal, postinfectious) glomerulonephritis, including but not limited to, poststreptococcal glomerulonephritis and nonstreptococcal acute glomerulonephritis, rapidly progressive (crescentic) glomerulonephritis, nephrotic syndrome, membranous glomerulonephritis (membranous nephropathy), minimal change disease (lipoid nephrosis), focal segmental glomerulosclerosis, membranoproliferative glomerulonephritis, IgA nephropathy (Berger disease), focal proliferative and necrotizing glomerulonephritis (focal glomerulonephritis), hereditary nephritis, including but not limited to, Alport syndrome and thin membrane disease (benign familial hematuria), chronic glomerulonephritis, glomerular lesions associated with systemic disease, including but not limited to, systemic lupus erythematosus., Henoch-Schönlein purpura, bacterial endocarditis, diabetic glomerulosclerosis, amyloidosis, fibrillary and immunotactoid glomerulonephritis, and other systemic disorders; diseases affecting tubules and interstitium, including acute tubular necrosis and tubulointerstitial nephritis, including but not limited to, pyelonephritis and urinary tract infection, acute pyelonephritis, chronic pyelonephritis and reflux nephropathy, and tubulointerstitial nephritis induced by drugs and toxins, including but not limited to, acute drug-induced interstitial nephritis, analgesic

abuse nephropathy, nephropathy associated with nonsteroidal anti-inflammatory drugs, and other tubulointerstitial diseases including, but not limited to, urate nephropathy, hypercalcemia and nephrocalcinosis, and multiple myeloma; diseases of blood vessels including benign nephrosclerosis, malignant hypertension and accelerated nephrosclerosis, renal artery stenosis, and thrombotic microangiopathies including, but not limited to, classic (childhood) hemolytic-uremic syndrome, adult hemolytic-uremic syndrome/thrombotic thrombocytopenic purpura, idiopathic HUS/TTP, and other vascular disorders including, but not limited to, atherosclerotic ischemic renal disease, atheroembolic renal disease, sickle cell disease nephropathy, diffuse cortical necrosis, and renal infarcts; urinary tract obstruction (obstructive uropathy); urolithiasis (renal calculi, stones); and tumors of the kidney including, but not limited to, benign tumors, such as renal papillary adenoma, renal fibroma or hamartoma (renomedullary interstitial cell tumor), angiomyolipoma, and oncocytoma, and malignant tumors, including renal cell carcinoma (hypemephroma, adenocarcinoma of kidney), which includes urothelial carcinomas of renal pelvis.

[0497] Examples of disorders of the lung include, but are not limited to, congenital anomalies; atelectasis; diseases of vascular origin, such as pulmonary congestion and edema, including hemodynamic pulmonary edema and edema caused by microvascular injury, adult respiratory distress syndrome (diffuse alveolar damage), pulmonary embolism, hemorrhage, and infarction, and pulmonary hypertension and vascular sclerosis; chronic obstructive pulmonary disease, such as emphysema, chronic bronchitis, bronchial asthma, and bronchiectasis; diffuse interstitial (infiltrative, restrictive) diseases, such as pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia (pulmonary infiltration with eosinophilia), *Bronchiolitis obliterans*-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, including Goodpasture syndrome, idiopathic pulmonary hemosiderosis and other hemorrhagic syndromes, pulmonary involvement in collagen vascular disorders, and pulmonary alveolar proteinosis; complications of therapies, such as drug-induced lung disease, radiation-induced lung disease, and lung transplantation; tumors, such as bronchogenic carcinoma, including paraneoplastic syndromes, bronchioloalveolar carcinoma, neuroendocrine tumors, such as bronchial carcinoid, miscellaneous tumors, and metastatic tumors; pathologies of the pleura, including inflammatory pleural effusions, noninflammatory pleural effusions, pneumothorax, and pleural tumors, including solitary fibrous tumors (pleural fibroma) and malignant mesothelioma.

[0498] As used herein, disorders involving the pancreas include those of the exocrine pancreas such as congenital anomalies, including but not limited to, ectopic pancreas; pancreatitis, including but not limited to, acute pancreatitis; cysts, including but not limited to, pseudocysts; tumors, including but not limited to, cystic tumors and carcinoma of the pancreas; and disorders of the endocrine pancreas such as, diabetes mellitus; islet cell tumors, including but not limited to, insulinomas, gastrinomas, and other rare islet cell tumors.

[0499] As used herein, disorders involving the ovary include, for example, polycystic ovarian disease, Stein-leventhal syndrome, Pseudomyxoma peritonei and stromal hyperthecosis; ovarian tumors such as, tumors of coelomic epithelium, serous tumors, mucinous tumors, endometeriod tumors, clear cell adenocarcinoma, cystadenofibroma, brenner tumor, surface epithelial tumors; germ cell tumors such as mature (benign) teratomas, monodermal teratomas, immature malignant teratomas, dysgerminoma, endodermal sinus tumor, choriocarcinoma; sex cord-stomal tumors such as, granulosa-theca cell tumors, thecoma-fibromas, androblastomas, hill cell tumors, and gonadoblastoma; and metastatic tumors such as Krukenberg tumors.

[0500] As used herein, hormonal disorders and diseases include type I and type II diabetes mellitus, pituitary disorders (e.g., growth disorders), thyroid disorders (e.g., hypothyroidism or hyperthyroidism), and reproductive or fertility disorders (e.g., disorders which affect the organs of the reproductive system, e.g., the prostate gland, the uterus, or the vagina; disorders which involve an imbalance in the levels of a reproductive hormone in a subject; disorders affecting the ability of a subject to reproduce; and disorders affecting secondary sex characteristic development, e.g., adrenal hyperplasia).

[0501] Aberrant expression and/or activity of the molecules of the invention can mediate disorders associated with bone metabolism. "Bone metabolism" refers to direct or indirect effects in the formation or degeneration of bone structures, e.g., bone formation, bone resorption, etc., which can ultimately affect the concentrations in serum of calcium and phosphate. This term also includes activities mediated by the molecules of the invention in bone cells, e.g. osteoclasts and osteoblasts, that can in turn result in bone formation and degeneration. For example, molecules of the invention can support different activities of bone resorbing osteoclasts such as the stimulation of differentiation of monocytes and mononuclear phagocytes into osteoclasts. Accordingly, molecules of the invention that modulate the production of bone cells can influence bone formation and degeneration, and thus can be used to treat bone disorders. Examples of such disorders include, but are not limited to, osteoporosis, osteodystrophy, osteomalacia, rickets, osteitis fibrosa cystica, renal osteodystrophy, osteosclerosis, anti-convulsant treatment, osteopenia, fibrogenesis-imperfecta ossium, secondary hyperparathyrodism, hypoparathyroidism, hyperparathyroidism, cirrhosis, obstructive jaundice, drug induced metabolism, medullary carcinoma, chronic renal disease, rickets, sarcoidosis, glucocorticoid antagonism, malabsorption syndrome, steatorrhea, tropical sprue, idiopathic hypercalcemia and milk fever.

[0502] As used herein, "a prostate disorder" refers to an abnormal condition occurring in the male pelvic region characterized by, e.g., male sexual dysfunction and/or urinary symptoms. This disorder may be manifested in the form of genitourinary inflammation (e.g., inflammation of smooth muscle cells) as in several common diseases of the prostate including prostatitis, benign prostatic hyperplasia and cancer, e.g., adenocarcinoma or carcinoma, of the prostate.

[0503] Examples of immune, e.g., inflammatory, (e.g. respiratory inflammatory) disorders or diseases include, but are not limited to, autoimmune diseases (including, for example, diabetes mellitus, arthritis (including rheumatoid arthritis, juvenile rheumatoid arthritis, osteoarthritis, psoriatic arthritis), multiple sclerosis, encephalomyelitis, myasthenia gravis, systemic lupus erythematosis, autoimmune thyroiditis, dermatitis (including atopic dermatitis and eczematous dermatitis), psoriasis, Sjögren's Syndrome, inflammatory bowel disease, e.g. Crohn's disease and ulcerative colitis, aphthous ulcer, iritis, conjunctivitis, keratoconjunctivitis, asthma, allergic asthma, chronic obstructive pulmonary disease, cutaneous lupus erythematosus, scleroderma, vaginitis, proctitis, drug eruptions, leprosy reversal reactions, erythema

nodosum leprosum, autoimmune uveitis, allergic encephalomyelitis, acute necrotizing hemorrhagic encephalopathy, idiopathic bilateral progressive sensorineural hearing loss, aplastic anemia, pure red cell anemia, idiopathic thrombocytopenia, polychondritis, Wegener's granulomatosis, chronic active hepatitis, Stevens-Johnson syndrome, idiopathic sprue, lichen planus, Graves' disease, sarcoidosis, primary biliary cirrhosis, uveitis posterior, and interstitial lung fibrosis), graft-versus-host disease, cases of transplantation, and allergy such as, atopic allergy.

[0504] As used herein, disorders involving the heart, or "cardiovascular disease" or a "cardiovascular disorder" includes a disease or disorder which affects the cardiovascular system, e.g., the heart, the blood vessels, and/or the blood. A cardiovascular disorder can be caused by an imbalance in arterial pressure, a malfunction of the heart, or an occlusion of a blood vessel, e.g., by a thrombus. A cardiovascular disorder includes, but is not limited to disorders such as arteriosclerosis, atherosclerosis, cardiac hypertrophy, ischemia reperfusion injury, restenosis, arterial inflammation, vascular wall remodeling, ventricular remodeling, rapid ventricular pacing, coronary microembolism, tachycardia, bradycardia, pressure overload, aortic bending, coronary artery ligation, vascular heart disease, valvular disease, including but not limited to, valvular degeneration caused by calcification, rheumatic heart disease, endocarditis, or complications of artificial valves; atrial fibrillation, long-QT syndrome, congestive heart failure, sinus node dysfunction, angina, heart failure, hypertension, atrial fibrillation, atrial flutter, pericardial disease, including but not limited to, pericardial effusion and pericarditis; cardiomyopathies, e.g., dilated cardiomyopathy or idiopathic cardiomyopathy, myocardial infarction, coronary artery disease, coronary artery spasm, ischemic disease, arrhythmia, sudden cardiac death, and cardiovascular developmental disorders (e.g., arteriovenous malformations, arteriovenous fistulae, raynaud's syndrome, neurogenic thoracic outlet syndrome, causalgia/reflex sympathetic dystrophy, hemangioma, aneurysm, cavernous angioma, aortic valve stenosis, atrial septal defects, atrioventricular canal, coarctation of the aorta, ebsteins anomaly, hypoplastic left heart syndrome, interruption of the aortic arch, mitral valve prolapse, ductus arteriosus, patent foramen ovale, partial anomalous pulmonary venous return, pulmonary atresia with ventricular septal defect, pulmonary atresia without ventricular septal defect, persistance of the fetal circulation, pulmonary valve stenosis, single ventricle, total anomalous pulmonary venous return, transposition of the great vessels, tricuspid atresia, truncus arteriosus, ventricular septal defects). A cardiovascular disease or disorder also can include an endothelial cell disorder.

[0505] As used herein, disorders involving the brain include, but are not limited to, disorders involving neurons, and disorders involving glia, such as astrocytes, oligodendrocytes, ependymal cells, and microglia; cerebral edema, raised intracranial pressure and herniation, and hydrocephalus; malformations and developmental diseases, such as neural tube defects, forebrain anomalies, posterior fossa anomalies, and syringomyelia and hydromyelia; perinatal brain injury; cerebrovascular diseases, such as those related to hypoxia, ischemia, and infarction, including hypotension, hypoperfusion, and low-flow states—global cerebral ischemia and focal cerebral ischemia—infarction from obstruction of local blood supply, intracranial hemorrhage, including intracerebral (intraparenchymal) hemorrhage, subarachnoid hemorrhage and ruptured berry aneurysms, and vascular malformations, hypertensive cerebrovascular disease, including lacunar infarcts, slit hemorrhages, and hypertensive encephalopathy; infections, such as acute meningitis, including acute pyogenic (bacterial) meningitis and acute aseptic (viral) meningitis, acute focal suppurative infections, including brain abscess, subdural empyema, and extradural abscess, chronic bacterial meningoencephalitis, including tuberculosis and mycobacterioses, neurosyphilis, and neuroborreliosis (Lyme disease), viral meningoencephalitis, including arthropod-borne (Arbo) viral encephalitis, Herpes simplex virus Type 1, Herpes simplex virus Type 2, Varicella-zoster virus (Herpes zoster), cytomegalovirus, poliomyelitis, rabies, and human immunodeficiency virus 1, including HIV-1 meningoencephalitis (subacute encephalitis), vacuolar myelopathy, AIDS-associated myopathy, peripheral neuropathy, and AIDS in children, progressive multifocal leukoencephalopathy, subacute sclerosing panencephalitis, fungal meningoencephalitis, other infectious diseases of the nervous system; transmissible spongiform encephalopathies (prion diseases); demyelinating diseases, including multiple sclerosis, multiple sclerosis variants, acute disseminated encephalomyelitis and acute necrotizing hemorrhagic encephalomyelitis, and other diseases with demyelination; degenerative diseases, such as degenerative diseases affecting the cerebral cortex, including Alzheimer disease and Pick disease, degenerative diseases of basal ganglia and brain stem, including Parkinsonism, idiopathic Parkinson disease (paralysis agitans), progressive supranuclear palsy, corticobasal degenration, multiple system atrophy, including striatonigral degenration, Shy-Drager syndrome, and olivopontocerebellar atrophy, and Huntington disease; spinocerebellar degenerations, including spinocerebellar ataxias, including Friedreich ataxia, and ataxia-telanglectasia, degenerative diseases affecting motor neurons, including amyotrophic lateral sclerosis (motor neuron disease), bulbospinal atrophy (Kennedy syndrome), and spinal muscular atrophy; inborn errors of metabolism, such as leukodystrophies, including Krabbe disease, metachromatic leukodystrophy, adrenoleukodystrophy, Pelizaeus-Merzbacher disease, and Canavan disease, mitochondrial encephalomyopathies, including Leigh disease and other mitochondrial encephalomyopathies; toxic and acquired metabolic diseases, including vitamin deficiencies such as thiamine (vitamin B_1) deficiency and vitamin B_{12} deficiency, neurologic sequelae of metabolic disturbances, including hypoglycemia, hyperglycemia, and hepatic encephatopathy, toxic disorders, including carbon monoxide, methanol, ethanol, and radiation, including combined methotrexate and radiation-induced injury; tumors, such as gliomas, including astrocytoma, including fibrillary (diffuse) astrocytoma and glioblastoma multiforme, pilocytic astrocytoma, pleomorphic xanthoastrocytoma, and brain stem glioma, oligodendroglioma, and ependymoma and related paraventricular mass lesions, neuronal tumors, poorly differentiated neoplasms, including medulloblastoma, other parenchymal tumors, including primary brain lymphoma, germ cell tumors, and pineal parenchymal tumors, meningiomas, metastatic tumors, paraneoplastic syndromes, peripheral nerve sheath tumors, including schwannoma, neurofibroma, and malignant peripheral nerve sheath tumor (malignant schwannoma), and neurocutaneous syndromes (phakomatoses), including neurofibromotosis, including Type 1 neurofibromatosis (NF1) and TYPE 2 neurofibromatosis (NF2), tuberous sclerosis, and Von Hippel-Lindau disease.

[0506] As used herein, disorders involving blood vessels include, but are not limited to, responses of vascular cell walls to injury, such as endothelial dysfunction and endothelial activation and intimal thickening; vascular diseases including, but not limited to, congenital anomalies, such as arteriovenous fistula, atherosclerosis, and hypertensive vascular disease, such as hypertension; inflammatory disease—the vasculitides, such as giant cell (temporal) arteritis, Takayasu arteritis, polyarteritis nodosa (classic), Kawasaki syndrome (mucocutaneous lymph node syndrome), microscopic polyanglitis (microscopic polyarteritis, hypersensitivity or leukocytoclastic anglitis), Wegener granulomatosis, thromboanglitis obliterans (Buerger disease), vasculitis associated with other disorders, and infectious arteritis; Raynaud disease; aneurysms and dissection, such as abdominal aortic aneurysms, syphilitic (luetic) aneurysms, and aortic dissection (dissecting hematoma); disorders of veins and lymphatics, such as varicose veins, thrombophlebitis and phlebothrombosis, obstruction of superior vena cava (superior vena cava syndrome), obstruction of inferior vena cava (inferior vena cava syndrome), and lymphangitis and lymphedema; tumors, including benign tumors and tumor-like conditions, such as hemangioma, lymphangioma, glomus tumor (glomangioma), vascular ectasias, and bacillary angiomatosis, and intermediate-grade (borderline low-grade malignant) tumors, such as Kaposi sarcoma and hemangloendothelioma, and malignant tumors, such as angiosarcoma and hemangiopericytoma; and pathology of therapeutic interventions in vascular disease, such as balloon angioplasty and related techniques and vascular replacement, such as coronary artery bypass graft surgery.

[0507] As used herein, disorders involving the testis and epididymis include, but are not limited to, congenital anomalies such as cryptorchidism, regressive changes such as atrophy, inflammations such as nonspecific epididymitis and orchitis, granulomatous (autoimmune) orchitis, and specific inflammations including, but not limited to, gonorrhea, mumps, tuberculosis, and syphilis, vascular disturbances including torsion, testicular tumors including germ cell tumors that include, but are not limited to, seminoma, spermatocytic seminoma, embryonal carcinoma, yolk sac tumor choriocarcinoma, teratoma, and mixed tumors, tumore of sex cord-gonadal stroma including, but not limited to, Leydig (interstitial) cell tumors and sertoli cell tumors (androblastoma), and testicular lymphoma, and miscellaneous lesions of tunica vaginalis.

[0508] As used herein, skeletal muscle, or musculoskeletal, disorders include, but are not limited to, muscular dystrophy (e.g., Duchenne muscular dystrophy, Becker muscular dystrophy, Emery-Dreifuss muscular dystrophy, limb-girdle muscular dystrophy, facioscapulohumeral muscular dystrophy, myotonic dystrophy, oculopharyngeal muscular dystrophy, distal muscular dystrophy, and congenital muscular dystrophy), motor neuron diseases (e.g., amyotrophic lateral sclerosis, infantile progressive spinal muscular atrophy, intermediate spinal muscular atrophy, spinal bulbar muscular atrophy, and adult spinal muscular atrophy), myopathies (e.g., inflammatory myopathies (e.g., dermatomyositis and polymyositis), myotonia congenita, paramyotonia congenita, central core disease, nemaline myopathy, myotubular myopathy, and periodic paralysis), tumors such as rhabdomyosarcoma, and metabolic diseases of muscle (e.g., phosphorylase deficiency, acid maltase deficiency, phosphofructokinase deficiency, debrancher enzyme deficiency, mitochondrial myopathy, carnitine deficiency, carnitine palmityl transferase deficiency, phosphoglycerate kinase deficiency, phosphoglycerate mutase deficiency, lactate dehydrogenase deficiency, and myoadenylate deaminase deficiency).

[0509] As used herein, an "endothelial cell disorder" includes a disorder characterized by aberrant, unregulated, or unwanted endothelial cell activity, e.g., proliferation, migration, angiogenesis, or vascularization; or aberrant expression of cell surface adhesion molecules or genes associated with angiogenesis, e.g., TIE-2, FLT and FLK. Endothelial cell disorders include tumorigenesis, tumor metastasis, psoriasis, diabetic retinopathy, endometriosis, Grave's disease, ischemic disease (e.g., atherosclerosis), and chronic inflammatory diseases (e.g., rheumatoid arthritis).

[0510] Disorders which can be treated or diagnosed by methods described herein include, but are not limited to, disorders associated with an accumulation in the liver of fibrous tissue, such as that resulting from an imbalance between production and degradation of the extracellular matrix accompanied by the collapse and condensation of preexisting fibers. The methods described herein can be used to diagnose or treat hepatocellular necrosis or injury induced by a wide variety of agents including processes which disturb homeostasis, such as an inflammatory process, tissue damage resulting from toxic injury or altered hepatic blood flow, and infections (e.g., bacterial, viral and parasitic). For example, the methods can be used for the early detection of hepatic injury, such as portal hypertension or hepatic fibrosis. In addition, the methods can be employed to detect liver fibrosis attributed to inborn errors of metabolism, for example, fibrosis resulting from a storage disorder such as Gaucher's disease (lipid abnormalities) or a glycogen storage disease, A1-antitrypsin deficiency; a disorder mediating the accumulation (e.g., storage) of an exogenous substance, for example, hemochromatosis (iron-overload syndrome) and copper storage diseases (Wilson's disease), disorders resulting in the accumulation of a toxic metabolite (e.g., tyrosinemia, fructosemia and galactosemia) and peroxisomal disorders (e.g., Zellweger syndrome). Additionally, the methods described herein can be used for the early detection and treatment of liver injury associated with the administration of various chemicals or drugs, such as for example, methotrexate, isonizaid, oxyphenisatin, methyldopa, chlorpromazine, tolbutamide or alcohol, or which represents a hepatic manifestation of a vascular disorder such as obstruction of either the intrahepatic or extrahepatic bile flow or an alteration in hepatic circulation resulting, for example, from chronic heart failure, veno-occlusive disease, portal vein thrombosis or Budd-Chiari syndrome.

[0511] As used herein, a "hepatic disorder" includes a disorder, disease or condition which affects the liver. The term hepatic disorder includes a disorder caused by the over- or under-production of hepatic enzymes, e.g., alanine aminotransferase, aspartate aminotransferase, or γ-glutammyl transferase, in the liver. For example, a hepatic disorder includes hepatic fibrosis, a hepatic disorder caused by a drug, a hepatic disorder caused by prolonged ethanol uptake, a hepatic injury caused by carbon tetrachloride exposure, hepatitis, liver tumors, cirrhosis of the liver, hemochromatosis, liver parasite induced disorders, alpha-1 antitrypsin deficiency, or autoimmune hepatitis. Hepatic disorders are disclosed at, for example, the American Liver Foundation website.

[0512] A hepatic disorder also includes a hepatic cell disorder. As used herein a "hepatic cell disorder" includes a disorder characterized by aberrant or unwanted hepatic cell activity, e.g., proliferation, migration, angiogenesis, or aberrant expression of cell surface adhesion molecules.

[0513] Additionally, the molecules of the invention can play an important role in the etiology of certain viral diseases, including but not limited to Hepatitis B, Hepatitis C and Herpes Simplex Virus (HSV). Modulators of the activity of the molecules of the invention could be used to control viral diseases. The modulators can be used in the treatment and/or diagnosis of viral infected tissue or virus-associated tissue fibrosis, especially liver and liver fibrosis. Also, such modulators can be used in the treatment and/or diagnosis of virus-associated carcinoma, especially hepatocellular cancer.

[0514] As used herein, a "viral pathogen" or "viral pathogen disorder" includes respiratory viral pathogens and their associated disorders include, for example, adenovirus, resulting in upper and lower respiratory tract infections; conjuctivitis and diarrhea; echovirus, resulting in upper respiratory tract infections, pharyngitis and rash; rhinovirus, resulting in upper respiratory tract infections; cosackievirus, resulting in Pleurodynia, herpangia, hand-foot-mouth disease; coronavirus, resulting in upper respiratory tract infections; influenza A and B viruses, resulting in influenza; parainfluenza virus 1-4, resulting in upper and lower respiratory tract infections and croup; respiratory syncytial virus, resulting in bronchiolitis and pneumonia. Digestive viral pathogens and their associated disorders include, for example, mumps virus, resulting in mumps, pancreatitis, and orchitis; rotavirus, resulting in childhood diarrhea; Norwalk Agent, resulting in gastroenteritis; hepatitis A virus, resulting in acute viral hepatitis; hepatitis B virus, hepatitis D virus and hepatitis C virus, resulting in acute or chronic hepatitis; hepatitis E virus, resulting in enterically transmitted hepatitis. Systemic viral pathogens associated with disorders involving skin eruptions include, for example, measles virus, resulting in measles (rubeola); rubella virus, resulting in German measles (rubella); parvovirus, resulting in erythema infectiosum and aplastic anemia; varicella-zoster virus, resulting in chicken pox and shingles; herpes simplex virus 1-associated, resulting in cold sores; and herpes simplex virus 2, resulting in genital herpes. Systemic viral pathogens associated with hematopoietic disorders include, for example, cytomegalovirus, resulting in cytomegalic inclusion disease; Epstein-Barr virus, resulting in mononucleosis; HTLV-1, resulting in adult T-cell leukemia and tropical spastic paraparesis; HTLV-II; and HIV 1 and HIV 2, resulting in AIDS. Arboviral pathogens associated with hemorrhagic fevers include, for example, dengue virus 1-4, resulting in dengue and hemorrhagic fever; yellow fever virus, resulting in yellow fever; Colorado tick fever virus, resulting in Colorado tick fever; and regional hemorrhagic fever viruses, resulting in Bolivian, Argentinian, Lassa fever. Viral pathogens associated with warty growths and other hyperplasias include, for example, papillomavirus, resulting in condyloma and cervical carcinoma; and molluscum virus, resulting in molluscum contagiosum. Viral pathogens associated with central nervous system disorders include, for example, poliovirus, resulting in poliomyelitis; rabiesvirus, associated with rabies; JC virus, associated with progressive multifocal leukoencephalophathy; and arboviral encephalitis viruses, resulting in Eastern, Western, Venezuelan, St. Louis, or California group encephalitis. Viral pathogens associated with cancer include, for example, human papillomaviruses, implicated in the genesis of several cancers including squamous cell carcinoma of the cervix and anogenital region, oral cancer and laryngeal cancers; Epstein-Barr virus, implicated in pathogenesis of the African form of Burkitt lymphoma, B-cell lymphomas, Hodgkin disease, and nasopharyngeal carcinomas; hepatitis B virus, implicated in liver cancer; human T-cell leukemia virus type 1 (HTLV-1), associated with T-cell leukemia/lymphoma; and the Kaposi sarcoma herpesvirus (KSHV).

[0515] "Blood platelet disorders" include, but are not limited to, thrombocytopenia due to a reduced number of megakaryocytes in the bone marrow, for example, as a result of chemotherapy; invasive disorders, such as leukemia, idiopathic or drug- or toxin-induced aplasia of the marrow, or rare hereditary amegakaryocytic thrombocytopenias; ineffective thrombopoiesis, for example, as a result of megaloblastic anemia, alcohol toxicity, vitamin B12 or folate deficiency, myelodysplastic disorders, or rare hereditary disorders (e.g., Wiskott-Aldrich syndrome and May-hegglin anomaly); a reduction in platelet distribution, for example, as a result of cirrhosis, a splenic invasive disease (e.g., Gaucher's disease), or myelofibrosis with extramedullary myeloid metaplasia; increased platelet destruction, for example, as a result of removal of IgG-coated platelets by the mononuclear phagocytic system (e.g., idiopathic thrombocytopenic purpura (ITP), secondary immune thrombocytopenia (e.g., systemic lupus erythematosus, lymphoma, or chronic lymphocytic leukemia), drug-related immune thrombocytopenias (e.g., as with quinidine, aspirin, and heparin), post-transfusion purpura, and neonatal thrombocytopenia as a result of maternal platelet autoantibodies or maternal platelet alloantibodies). Also included are thrombocytopenia secondary to intravascular clotting and thrombin induced damage to platelets as a result of, for example, obstetric complications, metastatic tumors, severe gram-negative bacteremia, thrombotic thrombocytopenic purpura, or severe illness. Also included is dilutional thrombocytopenia, for example, due to massive hemorrhage. Blood platelet disorders also include, but are not limited to, essential thrombocytosis and thrombocytosis associated with, for example, splenectomy, acute or chronic inflammatory diseases, hemolytic anemia, carcinoma, Hodgkin's disease, lymphoproliferative disorders, and malignant lymphomas.

[0516] Disorders related to reduced platelet number, thrombocytopenia, include idiopathic thrombocytopenic purpura, including acute idiopathic thrombocytopenic purpura, drug-induced thrombocytopenia, HIV-associated thrombocytopenia, and thrombotic microangiopathies: thrombotic thrombocytopenic purpura and hemolytic-uremic syndrome.

[0517] As used herein, neurological, central nervous system or neurodegenerative disorders include disorders of the central nervous system (CNS) and the peripheral nervous system, e.g., cognitive and neurodegenerative disorders, Examples of neurological disorders include, but are not limited to, autonomic function disorders such as hypertension and sleep disorders, and neuropsychiatric disorders, such as depression, schizophrenia, schizoaffective disorder, Korsakoff's psychosis, alcoholism, anxiety disorders, or phobic disorders; learning or memory disorders, e.g., amnesia or age-related memory loss, attention deficit disorder, dysthymic disorder, major depressive disorder, mania, obsessive-compulsive disorder, psychoactive substance use disorders, anxiety, phobias, panic disorder, as well as bipolar affective disorder, e.g., severe bipolar affective (mood) disorder (BP-

1), and bipolar affective neurological disorders, e.g., migraine and obesity. Such neurological disorders include, for example, disorders involving neurons, and disorders involving glia, such as astrocytes, oligodendrocytes, ependymal cells, and microglia; cerebral edema, raised intracranial pressure and herniation, and hydrocephalus; malformations and developmental diseases, such as neural tube defects, forebrain anomalies, posterior fossa anomalies, and syringomyelia and hydromyelia; perinatal brain injury; cerebrovascular diseases, such as those related to hypoxia, ischemia, and infarction, including hypotension, hypoperfusion, and low-flow states—global cerebral ischemia and focal cerebral ischemia—infarction from obstruction of local blood supply, intracranial hemorrhage, including intracerebral (intraparenchymal) hemorrhage, subarachnoid hemorrhage and ruptured berry aneurysms, and vascular malformations, hypertensive cerebrovascular disease, including lacunar infarcts, slit hemorrhages, and hypertensive encephalopathy; infections, such as acute meningitis, including acute pyogenic (bacterial) meningitis and acute aseptic (viral) meningitis, acute focal suppurative infections, including brain abscess, subdural empyema, and extradural abscess, chronic bacterial meningoencephalitis, including tuberculosis and mycobacterioses, neurosyphilis, and neuroborreliosis (Lyme disease), viral meningoencephalitis, including arthropod-borne (Arbo) viral encephalitis, Herpes simplex virus Type 1, Herpes simplex virus Type 2, Varicella-zoster virus (Herpes zoster), cytomegalovirus, poliomyelitis, rabies, and human immunodeficiency virus 1, including HIV-1 meningoencephalitis (subacute encephalitis), vacuolar myelopathy, AIDS-associated myopathy, peripheral neuropathy, and AIDS in children, progressive multifocal leukoencephalopathy, subacute sclerosing panencephalitis, fungal meningoencephalitis, other infectious diseases of the nervous system; transmissible spongiform encephalopathies (prion diseases); demyelinating diseases, including multiple sclerosis, multiple sclerosis variants, acute disseminated encephalomyelitis and acute necrotizing hemorrhagic encephalomyelitis, and other diseases with demyelination; degenerative diseases, such as degenerative diseases affecting the cerebral cortex, including Alzheimer's disease and Pick's disease, degenerative diseases of basal ganglia and brain stem, including Parkinsonism, idiopathic Parkinson's disease (paralysis agitans) and other Lewy diffuse body diseases, progressive supranuclear palsy, corticobasal degeneration, multiple system atrophy, including striatonigral degeneration, Shy-Drager syndrome, and olivopontocerebellar atrophy, and Huntington's disease, senile dementia, Gilles de la Tourette's syndrome, epilepsy, and Jakob-Creutzfieldt disease; spinocerebellar degenerations, including spinocerebellar ataxias, including Friedreich ataxia, and ataxia-telanglectasia, degenerative diseases affecting motor neurons, including amyotrophic lateral sclerosis (motor neuron disease), bulbospinal atrophy (Kennedy syndrome), and spinal muscular atrophy; inborn errors of metabolism, such as leukodystrophies, including Krabbe disease, metachromatic leukodystrophy, adrenoleukodystrophy, Pelizaeus-Merzbacher disease, and Canavan disease, mitochondrial encephalomyopathies, including Leigh disease and other mitochondrial encephalomyopathies; toxic and acquired metabolic diseases, including vitamin deficiencies such as thiamine (vitamin B_1) deficiency and vitamin B_{12} deficiency, neurologic sequelae of metabolic disturbances, including hypoglycemia, hyperglycemia, and hepatic encephatopathy, toxic disorders, including carbon monoxide, methanol, ethanol, and radiation, including combined methotrexate and radiation-induced injury; tumors, such as gliomas, including astrocytoma, including fibrillary (diffuse) astrocytoma and glioblastoma multiforme, pilocytic astrocytoma, pleomorphic xanthoastrocytoma, and brain stem glioma, oligodendroglioma, and ependymoma and related paraventricular mass lesions, neuronal tumors, poorly differentiated neoplasms, including medulloblastoma, other parenchymal tumors, including primary brain lymphoma, germ cell tumors, and pineal parenchymal tumors, meningiomas, metastatic tumors, paraneoplastic syndromes, peripheral nerve sheath tumors, including schwannoma, neurofibroma, and malignant peripheral nerve sheath tumor (malignant schwannoma), and neurocutaneous syndromes (phakomatoses), including neurofibromotosis, including Type 1 neurofibromatosis (NF1) and TYPE 2 neurofibromatosis (NF2), tuberous sclerosis, and Von Hippel-Lindau disease. Further CNS-related disorders include, for example, those listed in the American Psychiatric Association's Diagnostic and Statistical manual of Mental Disorders (DSM), the most current version of which is incorporated herein by reference in its entirety.

[0518] Disorders involving the eye and vision include, but are not limited to, granulomatous uveitis, cataracts, trachoma, corneal dystrophies, e.g., granular dystrophy or lattice dystrophy, glaucomas, retrolental fibroplasia, diabetes mellitus, hypertensive and arteriosclerotic retinopathy, retinitis pigmentosa, macular degeneration, retinoblastoma, papillaedema, and optic neuritis.

[0519] Additionally, molecules of the invention can play an important role in the regulation of metabolism or pain disorders. Diseases of metabolic imbalance include, but are not limited to, obesity, anorexia nervosa, cachexia, lipid disorders, and diabetes. Examples of pain disorders include headache posttherapeutic neuralgia, diabetic neuropathy, postmastectomy pain syndrome, stump pain, reflex sympathetic dystrophy, trigeminal neuralgia, neuropathic pain, orofacial neuropathic pain, osteoarthritis, arthritis, e.g., rheumatoid arthritis, fibromyalgia syndrome, tension myalgia, Guillian-Barre syndrome, Meralgia paraesthetica, burning mouth syndrome, fibrocitis, myofascial pain syndrome, idiopathic pain disorder, temporomandibular joint syndrome, atypical odontalgia, loin pain, haematuria syndrome, non-cardiac chest pain, back pain, chronic nonspecific pain, pain associated with surgery, psychogenic pain, tooth pain, musculoskeletal pain disorder, chronic pelvic pain, nonorganic chronic headache, tension-type headache, cluster headache, migraine, complex regional pain syndrome, vaginismus, nerve trunk pain, somatoform pain disorder, cyclical mastalgia, chronic fatigue syndrome, multiple somatization syndrome, chronic pain disorder, cancer pain, somatization disorder, Syndrome X, facial pain, idiopathic pain disorder, posttraumatic rheumatic pain modulation disorder (fibrositis syndrome), hyperalgesia, and Tangier disease.

[0520] As used herein, the term "pain signaling mechanisms" includes the cellular mechanisms involved in the development and regulation of pain, e.g., pain elicited by noxious chemical, mechanical, or thermal stimuli, in a subject, e.g., a mammal such as a human. In mammals, the initial detection of noxious chemical, mechanical, or thermal stimuli, a process referred to as "nociception", occurs predominantly at the peripheral terminals of specialized, small diameter sensory neurons. These sensory neurons transmit

the information to the central nervous system, evoking a perception of pain or discomfort and initiating appropriate protective reflexes.

[0521] Various aspects of the invention are described in further detail below.

Isolated Nucleic Acid Molecules

[0522] In one aspect, the invention provides, an isolated or purified, nucleic acid molecule that encodes a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide described herein, e.g., a full length 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or a fragment thereof, e.g., a biologically active portion of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. Also included is a nucleic acid fragment suitable for use as a hybridization probe, which can be used, e.g., to identify a nucleic acid molecule encoding a polypeptide of the invention, 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA, and fragments suitable for use as primers, e.g., PCR primers for the amplification or mutation of nucleic acid molecules.

[0523] In one embodiment, an isolated nucleic acid molecule of the invention includes the nucleotide sequence shown in SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, or a portion of any of this nucleotide sequence. In one embodiment, the nucleic acid molecule includes sequences encoding the human 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein (i.e., "the coding region" of SEQ ID NO:1, 5, 13, 18, 21, 24, 27, 32, 35, 38, 53, 56, 60, 69, 72 or 81, as shown in SEQ ID NO:3, 7, 15, 20, 23, 26, 29, 34, 37, 40, 55, 58, 62, 71, 74 or 83, respectively), as well as 5' untranslated sequences and 3' untranslated sequences. Alternatively, the nucleic acid molecule can include only the coding region of SEQ ID NO:1, 5, 13, 18, 21, 24, 27, 32, 35, 38, 53, 56, 60, 69, 72 or 81 (e.g., SEQ ID NO:3, 7, 15, 20, 23, 26, 29, 34, 37, 40, 55, 58, 62, 71, 74 or 83) and, e.g., no flanking sequences which normally accompany the subject sequence. In another embodiment, the nucleic acid molecule encodes a sequence corresponding to a fragment of the protein corresponding to domains within SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82.

[0524] In another embodiment, an isolated nucleic acid molecule of the invention includes a nucleic acid molecule which is a complement of the nucleotide sequence shown in SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, or a portion of any of these nucleotide sequences. In other embodiments, the nucleic acid molecule of the invention is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83 such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, thereby forming a stable duplex.

[0525] In one embodiment, an isolated nucleic acid molecule of the present invention includes a nucleotide sequence which is at least about: 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more homologous to the entire length of the nucleotide sequence shown in SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, or a portion, preferably of the same length, of any of these nucleotide sequences.

18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 Nucleic Acid Fragments

[0526] A nucleic acid molecule of the invention can include only a portion of the nucleic acid sequence of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83. For example, such a nucleic acid molecule can include a fragment which can be used as a probe or primer or a fragment encoding a portion of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, e.g., an immunogenic or biologically active portion of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. A fragment can comprise those nucleotides of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, which encode a domain of human 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751. The nucleotide sequence determined from the cloning of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene allows for the generation of probes and primers designed for use in identifying and/or cloning other 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 family members, or fragments thereof, as well as 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 homologs, or fragments thereof, from other species.

[0527] In another embodiment, a nucleic acid includes a nucleotide sequence that includes part, or all, of the coding region and extends into either (or both) the 5' or 3' noncoding region. Other embodiments include a fragment which includes a nucleotide sequence encoding an amino acid fragment described herein. Nucleic acid fragments can encode a specific domain or site described herein or fragments thereof, particularly fragments thereof which are at least 100 amino acids in length. Fragments also include nucleic acid sequences corresponding to specific amino acid sequences described above or fragments thereof. Nucleic acid fragments should not to be construed as encompassing those fragments that may have been disclosed prior to the invention.

[0528] A nucleic acid fragment can include a sequence corresponding to a domain, region, or functional site described herein. A nucleic acid fragment can also include one or more domain, region, or functional site described herein. Thus, for example, a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid fragment can include a sequence corresponding to a domain, as described herein.

[0529] 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 probes and primers are provided. Typically a probe/primer is an isolated or purified oligonucleotide. The oligonucleotide typically includes a region of nucleotide sequence that hybridizes under stringent conditions to at least about 7, 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 or antisense sequence of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83.

[0530] In a preferred embodiment the nucleic acid is a probe which is at least 5 or 10, and less than 200, more preferably less than 100, or less than 50, base pairs in length. It should be identical, or differ by 1, or less than in 5 or 10 bases, from a sequence disclosed herein. If alignment is needed for this comparison the sequences should be aligned for maximum homology. "Looped" out sequences from deletions or insertions, or mismatches, are considered differences.

[0531] A probe or primer can be derived from the sense or anti-sense strand of a nucleic acid which encodes a domain identified in the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequences, as disclosed herein.

[0532] In another embodiment a set of primers is provided, e.g., primers suitable for use in a PCR, which can be used to amplify a selected region of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence, e.g., a domain, region, site or other sequence described herein. The primers should be at least 5, 10, or 50 base pairs in length and less than 100, or less than 200, base pairs in length. The primers should be identical, or differ by one base from a sequence disclosed herein or from a naturally occurring variant.

[0533] A nucleic acid fragment can encode an epitope bearing region of a polypeptide described herein.

[0534] A nucleic acid fragment encoding a "biologically active portion of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide" can be prepared by isolating a portion of the nucleotide sequence of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, which encodes a polypeptide having a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 biological activity (e.g., the biological activities of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins are described herein), expressing the encoded portion of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein (e.g., by recombinant expression in vitro) and assessing the activity of the encoded portion of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. A nucleic acid fragment encoding a biologically active portion of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide, can comprise a nucleotide sequence which is greater than 300 or more nucleotides in length.

[0535] In preferred embodiments, a nucleic acid includes a nucleotide sequence which is about 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, 3300, 3400, 3500, 3600, 3700, 3800, 3900, 4000, 4100, 4200, 4300, 4400, 4500 or more nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83.

18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 Nucleic Acid Variants

[0536] The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83. Such differences can be due to degeneracy of the genetic code (and result in a nucleic acid which encodes the same 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins as those encoded by the nucleotide sequence disclosed herein. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence which differs, by at least 1, but less than 5, 10, 20, 50, or 100 amino acid residues that shown in SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82. If alignment is needed for this comparison the sequences should be aligned for maximum homology. "Looped" out sequences from deletions or insertions, or mismatches, are considered differences.

[0537] Nucleic acids of the inventor can be chosen for having codons, which are preferred, or non-preferred, for a particular expression system. E.g., the nucleic acid can be one in which at least one codon, at preferably at least 10%, or 20% of the codons has been altered such that the sequence is optimized for expression in *E. coli*, yeast, human, insect, or CHO cells.

[0538] Nucleic acid variants can be naturally occurring, such as allelic variants (same locus), homologs (different locus), and orthologs (different organism) or can be non naturally occurring. Non-naturally occurring variants can be made by mutagenesis techniques, including those applied to polynucleotides, cells, or organisms. The variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions (as compared in the encoded product).

[0539] In a preferred embodiment, the nucleic acid differs from that of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, e.g., as follows: by at least one but less than 10, 20, 30, or 40 nucleotides; at least one but less than 1%, 5%, 10% or 20% of the nucleotides in the subject nucleic acid. If necessary for this analysis the sequences should be aligned for maximum homology. "Looped" out sequences from deletions or insertions, or mismatches, are considered differences.

[0540] Orthologs, homologs, and allelic variants can be identified using methods known in the art. These variants comprise a nucleotide sequence encoding a polypeptide that is 50%, at least about 55%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more identical to the nucleotide sequence shown in SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 or a fragment of this sequence. Such nucleic acid molecules can readily be identified as being able to hybridize under stringent conditions, to the nucleotide sequence shown in SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 or a fragment of the sequence. Nucleic acid molecules corresponding to orthologs,

homologs, and allelic variants of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 cDNAs of the invention can further be isolated by mapping to the same chromosome or locus as the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene.

[0541] Preferred variants include those that are correlated with activities specific to the molecules of the invention, i.e. calcium channel activity, calcium/sodium antiporter activity, potassium channel activity, organic ion transporter activity, choline transporter activity, or other.

[0542] Allelic variants of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751, e.g., human 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751, include both functional and non-functional proteins. Functional allelic variants are naturally occurring amino acid sequence variants of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein within a population that maintain the ability to bind a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 ligand or substrate and/or modulate calcium channel activity, calcium/sodium antiporter activity, potassium channel activity, organic ion transporter activity or choline transporter activity. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751, e.g., human 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751, protein within a population that do not have the ability to bind a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 ligand or substrate and/or modulate calcium channel activity, calcium/sodium antiporter activity, potassium channel activity, organic ion transporter activity or choline transporter activity. Non-functional allelic variants will typically contain a non-conservative substitution, a deletion, or insertion, or premature truncation of the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, or a substitution, insertion, or deletion in critical residues or critical regions of the protein.

[0543] Moreover, nucleic acid molecules encoding other 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 family members and, thus, which have a nucleotide sequence which differs from the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequences of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83 are intended to be within the scope of the invention.

Antisense Nucleic Acid Molecules, Ribozymes and Modified 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 Nucleic Acid Molecules

[0544] In another aspect, the invention features, an isolated nucleic acid molecule which is antisense to 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751. An "antisense" nucleic acid can include 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. The antisense nucleic acid can be complementary to an entire 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 coding strand, or to only a portion thereof (e.g., the coding region of human 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 corresponding to SEQ ID NO:3, 7, 15, 20, 23, 26, 29, 34, 37, 40, 55, 58, 62, 71, 74 or 83, respectively). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 (e.g., the 5' and 3' untranslated regions).

[0545] An antisense nucleic acid can be designed such that it is complementary to the entire coding region of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA, e.g., between the −10 and +10 regions of the target gene nucleotide sequence of interest. An antisense oligonucleotide can be, for example, about 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or more nucleotides in length.

[0546] 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. The antisense nucleic acid also 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, described further in the following subsection).

[0547] The antisense nucleic acid molecules of the invention are typically administered to a subject (e.g., by direct injection at a tissue site), or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For systemic administration, antisense molecules can be modified such that they specifically or selectively 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.

[0548] In yet another embodiment, the antisense nucleic acid molecule 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).

[0549] In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. A ribozyme having specificity for a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-encoding nucleic acid can include one or more sequences complementary to the nucleotide sequence of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 cDNA disclosed herein (i.e., SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83), and a sequence having known catalytic sequence responsible for mRNA cleavage (see U.S. Pat. No. 5,093,246 or Haselhoff and Gerlach (1988) *Nature* 334:585-591). 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-encoding mRNA. See, e.g., Cech et al. U.S. Pat. No. 4,987,071; and Cech et al. U.S. Pat. No. 5,116,742. Alternatively, 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel and Szostak (1993) *Science* 261:1411-1418.

[0550] 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 (e.g., the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 promoter and/or enhancers) to form triple helical structures that prevent transcription of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene in target cells. See generally, Helene (1991) Anticancer Drug Des. 6:569-84; Helene (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14:807-15. The potential sequences that can be targeted for triple helix formation can be increased by creating a so-called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

[0551] The invention also provides detectably labeled oligonucleotide primer and probe molecules. Typically, such labels are chemiluminescent, fluorescent, radioactive, or colorimetric.

[0552] A 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid molecule 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 et al. (1996) *Bioorganic & Medicinal Chemistry* 4: 5-23).

[0553] As used herein, the terms "peptide nucleic acid" or "PNA" refers to a nucleic acid mimic, e.g., a DNA mimic, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of a PNA can 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 et al. (1996) supra; Perry-O'Keefe et al. (1996) *Proc. Natl. Acad. Sci.* 93: 14670-675.

[0554] PNAs of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid molecules can be used in therapeutic and diagnostic applications. 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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 et al. (1996) supra)); or as probes or primers for DNA sequencing or hybridization (Hyrup et al. (1996) supra; Perry-O'Keefe supra).

[0555] In other embodiments, the oligonucleotide can 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. WO88/09810) or the blood-brain barrier (see, e.g., PCT Publication No. WO89/10134). 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 can be conjugated to another molecule, (e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

[0556] The invention also includes molecular beacon oligonucleotide primer and probe molecules having at least one region which is complementary to a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid of the invention, two complementary regions one having a fluorophore and one a quencher such that the molecular beacon is useful for quan-

titating the presence of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid of the invention in a sample. Molecular beacon nucleic acids are described, for example, in Lizardi et al., U.S. Pat. No. 5,854,033; Nazarenko et al., U.S. Pat. No. 5,866,336, and Livak et al., U.S. Pat. No. 5,876,930.

Isolated 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 Polypeptides

[0557] In another aspect, the invention features, an isolated 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, or fragment, e.g., a biologically active portion, for use as immunogens or antigens to raise or test (or more generally to bind) anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibodies. 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein can be isolated from cells or tissue sources using standard protein purification techniques. 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or fragments thereof can be produced by recombinant DNA techniques or synthesized chemically.

[0558] Polypeptides of the invention include those which arise as a result of the existence of multiple genes, alternative transcription events, alternative RNA splicing events, and alternative translational and post-translational events. The polypeptide can be expressed in systems, e.g., cultured cells, which result in substantially the same post-translational modifications present when the polypeptide is expressed in a native cell, or in systems which result in the alteration or omission of post-translational modifications, e.g., glycosylation or cleavage, present in a native cell.

[0559] In a preferred embodiment, a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide has one or more of the following characteristics: 1) it has the ability to modulate membrane excitability; 2) it has the ability to influence the resting potential of membranes; 3) it has the ability to modulate wave forms and frequencies of action potentials; 4) it has the ability to modulate thresholds of excitation; 5) it has the ability to modulate neurite outgrowth and synaptogenesis; 6) it had the ability to modulate signal transduction, 7) it has the ability to bind a second messenger; 8) it has the ability to bind diacylglycerol; 9) it has the ability to regulate the flow of cations through a membrane; 10) it has the ability to transport a substrate or target molecule, e.g., an ion (e.g., a calcium ion) across a membrane; 11) it has the ability to transport a second substrate or target molecule, e.g., another ion (e.g., a sodium ion) across a membrane; 12) it has the ability to transport a third substrate or target molecule, e.g., another ion (e.g., a potassium ion) across a membrane; 13) it has the ability to interact with and/or modulate the activity of a second non-transporter protein; 14) it has the ability to modulate cellular signaling and/or gene transcription (e.g., either directly or indirectly; 15) it has the ability to interact with a non-TWIK protein molecule; 16) it has the ability to activate a TWIK-dependent signal transduction pathway; 17) it has the ability to modulate the release of neurotransmitters; 18) it has the ability to protect cells and/or tissues from organic ions; 19) it has the ability to modulate intracellular Ca^{2+} concentration; 20) it has the ability to bind a ligand, e.g., L-glutamate, and/or glycine; 21) it has the ability to modulate (e.g., promote, catalyze, regulate, initiate, facilitate or inhibit) the manufacture of choline metabolites and/or compounds of which choline is a component or precursor, e.g., phospholipids (e.g., phosphatidylcholine (lecithin), sphingomyelin, sphingophosphorylcholine, and platelet activating factor), acetylcholine, very low density lipoproteins (VLDLs), and betaine, e.g., by transporting choline into or out of cells; 22) it has the ability to modulate (e.g., promote, catalyze, regulate, initiate, facilitate or inhibit) transport of choline, its metabolites, and/or compounds of which choline is a component or precursor across membranes (e.g., plasma membranes), e.g., from an extracellular medium into a cell, or vice versa; 23) it has the ability to modulate (e.g., promote, catalyze, regulate, initiate, facilitate or inhibit) transport of choline, its metabolites, and/or compounds of which choline is a component or precursor across barriers between tissues (e.g., the blood-brain barrier); 24) it has a molecular weight, e.g., a deduced molecular weight, preferably ignoring any contribution of post translational modifications, amino acid composition or other physical characteristic of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide, e.g., a polypeptide of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82; 25) it has an overall sequence similarity of at least 60%, preferably at least 70%, more preferably at least 80, 90, or 95%, with a polypeptide of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82; 26) it is expressed in a multitude of human tissues and cell lines (refer to section for each molecule of the invention); and 27) it has specific domains which are preferably about 70%, 80%, 90% or 95% identical to the identified amino acid residues of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 (refer to section for each molecule of the invention for domain names and locations within amino acid sequence).

[0560] In a preferred embodiment the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, or fragment thereof, differs from the corresponding sequence in SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82. In one embodiment it differs by at least one but by less than 15, 10 or 5 amino acid residues. In another it differs from the corresponding sequence in SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 by at least one residue but less than 20%, 15%, 10% or 5% of the residues in it differ from the corresponding sequence in SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82. (If this comparison requires alignment the sequences should be aligned for maximum homology. "Looped" out sequences from deletions or insertions, or mismatches, are considered differences.) The differences are, preferably, differences or changes at a non-essential residue or a conservative substitution. In a preferred embodiment the differences are not in the identified or conserved domain(s) within SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82. In another embodiment one or more differences are in the cidentified or conserved domain(s) within SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82.

[0561] Other embodiments include a protein that contains one or more changes in amino acid sequence, e.g., a change in an amino acid residue which is not essential for activity. Such 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 pro-

teins differ in amino acid sequence from SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, yet retain biological activity.

[0562] In one embodiment, the protein includes an amino acid sequence at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82.

[0563] A 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or fragment is provided which varies from the sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 in regions defined by amino acids that are not within identified or conserved domains or regions by at least one but by less than 15, 10 or 5 amino acid residues in the protein or fragment but which does not differ from SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 in regions defined by amino acids that are within identified or conserved domains or regions. (If this comparison requires alignment the sequences should be aligned for maximum homology. "Looped" out sequences from deletions or insertions, or mismatches, are considered differences.) In some embodiments the difference is at a non-essential residue or is a conservative substitution, while in others the difference is at an essential residue or is a non-conservative substitution.

[0564] In one embodiment, a biologically active portion of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein includes an identified domain (refer to section for each molecule of the invention). Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein.

[0565] In a preferred embodiment, the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein has an amino acid sequence shown in SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82. In other embodiments, the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein is sufficiently or substantially identical to SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82. In yet another embodiment, the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein is sufficiently or substantially identical to SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 and retains the functional activity of the protein of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, as described in detail in the subsections above.

18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 Chimeric or Fusion Proteins

[0566] In another aspect, the invention provides 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 chimeric or fusion proteins. As used herein, a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 "chimeric protein" or "fusion protein" includes a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide linked to a non-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide. A "non-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, e.g., a protein which is different from the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein and which is derived from the same or a different organism. The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide of the fusion protein can correspond to all or a portion e.g., a fragment described herein of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 amino acid sequence. In a preferred embodiment, a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 fusion protein includes at least one (or two) biologically active portion of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. The non-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide can be fused to the N-terminus or C-terminus of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide.

[0567] The fusion protein can include a moiety which has a high affinity for a ligand. For example, the fusion protein can be a GST-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 fusion protein in which the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751. Alternatively, the fusion protein can be a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression and/or secretion of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 can be increased through use of a heterologous signal sequence.

[0568] Fusion proteins can include all or a part of a serum protein, e.g., a portion of an immunoglobulin (e.g., IgG, IgA, or IgE), e.g., an Fc region and/or the hinge C1 and C2 sequences of an immunoglobulin or human serum albumin.

[0569] The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject in vivo. The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 fusion proteins can be used to affect the bioavailability of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 substrate. 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 fusion proteins can be useful therapeutically for the treatment

of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein; (ii) mis-regulation of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene; and (iii) aberrant post-translational modification of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein.

[0570] Moreover, the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-fusion proteins of the invention can be used as immunogens to produce anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibodies in a subject, to purify 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 ligands and in screening assays to identify molecules which inhibit the interaction of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 with a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 substrate.

[0571] Expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein.

Variants of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 Proteins

[0572] In another aspect, the invention also features a variant of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide, e.g., which functions as an agonist (mimetics) or as an antagonist. Variants of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins can be generated by mutagenesis, e.g., discrete point mutation, the insertion or deletion of sequences or the truncation of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. An agonist of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. An antagonist of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein can inhibit one or more of the activities of the naturally occurring form of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein by, for example, competitively modulating a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-mediated activity of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. Preferably, 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein.

[0573] Variants of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein for agonist or antagonist activity.

[0574] Libraries of fragments e.g., N terminal, C terminal, or internal fragments, of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein coding sequence can be used to generate a variegated population of fragments for screening and subsequent selection of variants of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein.

[0575] Variants in which a cysteine residues is added or deleted or in which a residue which is glycosylated is added or deleted are particularly preferred.

[0576] Methods 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 are known in the art. 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 variants (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave et al. (1993) *Protein Engineering* 6:327-331).

[0577] Cell based assays can be exploited to analyze a variegated 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 library. For example, a library of expression vectors can be transfected into a cell line, e.g., a cell line, which ordinarily responds to 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 in a substrate-dependent manner. The transfected cells are then contacted with 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 and the effect of the expression of the mutant on signaling by the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 substrate can be detected, e.g., by measuring either calcium channel activity, calcium/sodium antiporter activity, potassium channel activity, organic ion transporter activity, choline transporter activity, or other activity. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of signaling by the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 substrate, and the individual clones further characterized.

[0578] In another aspect, the invention features a method of making a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide, e.g., a peptide having a non-wild type activity, e.g., an antagonist, agonist, or super agonist of a naturally occurring 18607, 15603, 69318, 12303, 48000, 52920, 5433,

38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide, e.g., a naturally occurring 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide. The method includes altering the sequence of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide, e.g., altering the sequence, e.g., by substitution or deletion of one or more residues of a non-conserved region, a domain or residue disclosed herein, and testing the altered polypeptide for the desired activity.

[0579] In another aspect, the invention features a method of making a fragment or analog of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide a biological activity of a naturally occurring 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide. The method includes altering the sequence, e.g., by substitution or deletion of one or more residues, of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide, e.g., altering the sequence of a non-conserved region, or a domain or residue described herein, and testing the altered polypeptide for the desired activity.

Anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 Antibodies

[0580] In another aspect, the invention provides an anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibody. The term "antibody" as used herein refers to an immunoglobulin molecule or immunologically active portion thereof, i.e., an antigen-binding portion. Examples of immunologically active portions of immunoglobulin molecules include scFV and dcFV fragments, Fab and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as papain or pepsin, respectively.

[0581] The antibody can be a polyclonal, monoclonal, recombinant, e.g., a chimeric or humanized, fully human, non-human, e.g., murine, or single chain antibody. In a preferred embodiment it has effector function and can fix complement. The antibody can be coupled to a toxin or imaging agent.

[0582] A full-length 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or, antigenic peptide fragment of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 can be used as an immunogen or can be used to identify anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibodies made with other immunogens, e.g., cells, membrane preparations, and the like. The antigenic peptide of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 should include at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 and encompasses an epitope of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751. Preferably, the antigenic peptide includes 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.

[0583] Fragments of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 which include hydrophilic regions of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 can be used to make, e.g., used as immunogens or used to characterize the specificity of an antibody, antibodies against hydrophilic regions of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. Similarly, fragments of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 which include hydrophobic regions of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 can be used to make an antibody against a hydrophobic region of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein; fragments of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 which include residues within extra cellular domain(s) of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 can be used to make an antibody against an extracellular or non-cytoplasmic region of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein; fragments of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 which include residues within intracellular regions of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 can be used to make an antibody against an intracellular region of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein; a fragment of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 which include residues within identified or conserved domains of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82 can be used to make an antibody against the identified or conserved domain of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein.

[0584] Antibodies reactive with, or specific or selective for, any of these regions, or other regions or domains described herein are provided.

[0585] Preferred epitopes encompassed by the antigenic peptide are regions of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 located on the surface of the protein, e.g., hydrophilic regions, as well as regions with high antigenicity. For example, an Emini surface probability analysis of the human 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein sequence can be used to indicate the regions that have a particularly high probability of being localized to the surface of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein and are thus likely to constitute surface residues useful for targeting antibody production.

[0586] In a preferred embodiment the antibody can bind to the extracellular portion of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, e.g., it can bind to a whole cell which expresses the 18607, 15603, 69318, 12303, 48000,

52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. In another embodiment, the antibody binds an intracellular portion of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein.

[0587] In a preferred embodiment the antibody binds an epitope on any domain or region on 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins described herein.

[0588] Additionally, chimeric, humanized, and completely human antibodies are also within the scope of the invention. Chimeric, humanized, but most preferably, completely human antibodies are desirable for applications which include repeated administration, e.g., therapeutic treatment of human patients, and some diagnostic applications.

[0589] Chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, can be made using standard recombinant DNA techniques. 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, 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. Pat. 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; and Shaw et al. (1988) *J. Natl. Cancer Inst.* 80:1553-1559).

[0590] A humanized or complementarity determining region (CDR)-grafted antibody will have at least one or two, but generally all three recipient CDR's (of heavy and or light immuoglobulin chains) replaced with a donor CDR. The antibody may be replaced with at least a portion of a non-human CDR or only some of the CDR's may be replaced with non-human CDR's. It is only necessary to replace the number of CDR's required for binding of the humanized antibody to a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 or a fragment thereof. Preferably, the donor will be a rodent antibody, e.g., a rat or mouse antibody, and the recipient will be a human framework or a human consensus framework. Typically, the immunoglobulin providing the CDR's is called the "donor" and the immunoglobulin providing the framework is called the "acceptor." In one embodiment, the donor immunoglobulin is a non-human (e.g., rodent). The acceptor framework is a naturally-occurring (e.g., a human) framework or a consensus framework, or a sequence about 85% or higher, preferably 90%, 95%, 99% or higher identical thereto.

[0591] As used herein, the term "consensus sequence" refers to the sequence formed from the most frequently occurring amino acids (or nucleotides) in a family of related sequences (See e.g., Winnaker, (1987) *From Genes to Clones* (Verlagsgesellschaft, Weinheim, Germany). In a family of proteins, each position in the consensus sequence is occupied by the amino acid occurring most frequently at that position in the family. If two amino acids occur equally frequently, either can be included in the consensus sequence. A "consensus framework" refers to the framework region in the consensus immunoglobulin sequence.

[0592] An antibody can be humanized by methods known in the art. Humanized antibodies can be generated by replacing sequences of the Fv variable region which are not directly involved in antigen binding with equivalent sequences from human Fv variable regions. General methods for generating humanized antibodies are provided by Morrison (1985) *Science* 229:1202-1207, by Oi et al. (1986) *BioTechniques* 4:214, and by Queen et al. U.S. Pat. Nos. 5,585,089, 5,693,761 and 5,693,762, the contents of all of which are hereby incorporated by reference. Those methods include isolating, manipulating, and expressing the nucleic acid sequences that encode all or part of immunoglobulin Fv variable regions from at least one of a heavy or light chain. Sources of such nucleic acid are well known to those skilled in the art and, for example, may be obtained from a hybridoma producing an antibody against a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide or fragment thereof. The recombinant DNA encoding the humanized antibody, or fragment thereof, can then be cloned into an appropriate expression vector.

[0593] Humanized or CDR-grafted antibodies can be produced by CDR-grafting or CDR substitution, wherein one, two, or all CDR's of an immunoglobulin chain can be replaced. See e.g., U.S. Pat. No. 5,225,539; Jones et al. (1986) *Nature* 321:552-525; Verhoeyan et al. (1988) *Science* 239:1534; Beidler et al. (1988) *J. Immunol.* 141:4053-4060; Winter U.S. Pat. No. 5,225,539, the contents of all of which are hereby expressly incorporated by reference. Winter describes a CDR-grafting method which may be used to prepare the humanized antibodies of the present invention (UK Patent Application GB 2188638A, filed on Mar. 26, 1987; Winter U.S. Pat. No. 5,225,539), the contents of which is expressly incorporated by reference.

[0594] Also within the scope of the invention are humanized antibodies in which specific amino acids have been substituted, deleted or added. Preferred humanized antibodies have amino acid substitutions in the framework region, such as to improve binding to the antigen. For example, a humanized antibody will have framework residues identical to the donor framework residue or to another amino acid other than the recipient framework residue. To generate such antibodies, a selected, small number of acceptor framework residues of the humanized immunoglobulin chain can be replaced by the corresponding donor amino acids. Preferred locations of the substitutions include amino acid residues adjacent to the CDR, or which are capable of interacting with a CDR (see e.g., U.S. Pat. No. 5,585,089). Criteria for selecting amino acids from the donor are described in U.S. Pat. No. 5,585,089, e.g., columns 12-16 of U.S. Pat. No. 5,585,089, the e.g., columns 12-16 of U.S. Pat. No. 5,585,089, the contents of which are hereby incorporated by reference. Other techniques for humanizing antibodies are described in Padlan et al. EP 519596 A1, published on Dec. 23, 1992.

[0595] Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Such antibodies can be produced using transgenic mice that are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. See, for example, Lonberg and Huszar (1995) *Int. Rev. Immunol.* 13:65-93); and U.S. Pat. Nos. 5,625,126; 5,633,425; 5,569,825; 5,661,016; and 5,545,806. In addition, companies such as Abgenix, Inc. (Fremont, Calif.) and Medarex, Inc. (Princeton, N.J.), can be engaged to

provide human antibodies directed against a selected antigen using technology similar to that described above.

[0596] Completely human antibodies that recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, e.g., a murine antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. This technology is described by Jespers et al. (1994) *Bio/Technology* 12:899-903).

[0597] The anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibody can be a single chain antibody. A single-chain antibody (scFV) can be engineered as described in, for example, Colcher et al. (1999) *Ann. N Y Acad. Sci.* 880:263-80; and Reiter (1996) *Clin. Cancer Res.* 2:245-52. The single chain antibody can be dimerized or multimerized to generate multivalent antibodies having specificities for different epitopes of the same target 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein.

[0598] In a preferred embodiment, the antibody has reduced or no ability to bind an Fc receptor. For example, it is an isotype or subtype, fragment or other mutant, which does not support binding to an Fc receptor, e.g., it has a mutagenized or deleted Fc receptor binding region.

[0599] An antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive 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, puromycin, maytansinoids, e.g., maytansinol (see U.S. Pat. No. 5,208,020), CC-1065 (see U.S. Pat. Nos. 5,475,092, 5,585,499, 5,846,545) 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, CC-1065, 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, vinblastine, taxol and maytansinoids).

[0600] Radioactive ions include, but are not limited to iodine, yttrium and praseodymium.

[0601] The conjugates of the invention can be used for modifying a given biological response, the therapeutic moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the therapeutic 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, α-interferon, β-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.

[0602] Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Pat. No. 4,676,980.

[0603] An anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibody (e.g., monoclonal antibody) can be used to isolate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, an anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibody can be used to detect 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the protein. Anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance (i.e., antibody labelling). 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}U, ^{35}S or ^{3}H.

[0604] In preferred embodiments, an antibody can be made by immunizing with a purified 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antigen, or a fragment thereof, e.g., a fragment described herein, a membrane associated antigen, tissues, e.g., crude tissue preparations, whole cells, preferably living cells, lysed cells, or cell fractions, e.g., membrane fractions.

[0605] Antibodies which bind only a native 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, only denatured or otherwise non-native 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, or which bind both, are within the invention. Antibodies with linear or conformational epitopes are within the invention. Conformational epitopes sometimes can be identified by identifying antibodies which bind to native but not denatured 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein.

Recombinant Expression Vectors, Host Cells and Genetically Engineered Cells

[0606] In another aspect, the invention includes, vectors, preferably expression vectors, containing a nucleic acid

encoding a polypeptide described herein. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked and can include a plasmid, cosmid or viral vector. The vector can be capable of autonomous replication or it can integrate into a host DNA. Viral vectors include, e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses.

[0607] A vector can include a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid in a form suitable for expression of the nucleic acid in a host cell.

[0608] Preferably the recombinant expression vector includes one or more regulatory sequences operatively linked to the nucleic acid sequence to be expressed. The term "regulatory sequence" includes promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence, as well as tissue-specific regulatory and/or inducible sequences. 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 polypeptides, including fusion proteins or polypeptides, encoded by nucleic acids as described herein (e.g., 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins, mutant forms of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins, fusion proteins, and the like).

[0609] The recombinant expression vectors of the invention can be designed for expression of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins in prokaryotic or eukaryotic cells. For example, polypeptides of the invention can be expressed in *E. coli*, insect cells (e.g., using baculovirus expression vectors), yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, (1990) *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Calif. Alternatively, the recombinant expression vector can be transcribed and translated in vitro, for example using T7 promoter regulatory sequences and T7 polymerase.

[0610] 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, 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 and Johnson (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

[0611] Purified fusion proteins can be used in 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity assays, (e.g., direct assays or competitive assays described in detail below), or to generate antibodies specific or selective for 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins. In a preferred embodiment, a fusion protein expressed in a retroviral expression vector of the present invention can be used 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).

[0612] To maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman (1990) *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Calif. 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada et al., (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

[0613] The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression vector can be a yeast expression vector, a vector for expression in insect cells, e.g., a baculovirus expression vector or a vector suitable for expression in mammalian cells.

[0614] 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.

[0615] 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). Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert et al. (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) EMBO J. 8:729-733) and immunoglobulins (Banerji et al. (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund et al. (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example, the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

[0616] The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orien-

tation. Regulatory sequences (e.g., viral promoters and/or enhancers) operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the constitutive, tissue specific or cell type specific expression of antisense RNA in a variety of cell types. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus. For a discussion of the regulation of gene expression using antisense genes see Weintraub et al., (1986) *Reviews—Trends in Genetics* 1:1.

[0617] Another aspect the invention provides a host cell which includes a nucleic acid molecule described herein, e.g., a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid molecule within a recombinant expression vector or a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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. Such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications can 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.

[0618] A host cell can be any prokaryotic or eukaryotic cell. For example, a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary (CHO) cells or CV-1 origin, SV-40 (COS) cells). Other suitable host cells are known to those skilled in the art.

[0619] Vector DNA can be introduced into host 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.

[0620] A host cell of the invention can be used to produce (i.e., express) a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. Accordingly, the invention further provides methods for producing a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein using the host cells of the invention. In one embodiment, the method includes culturing the host cell of the invention (into which a recombinant expression vector encoding a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein has been introduced) in a suitable medium such that a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein is produced. In another embodiment, the method further includes isolating a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein from the medium or the host cell.

[0621] In another aspect, the invention features, a cell or purified preparation of cells which include a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 transgene, or which otherwise misexpress 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751. The cell preparation can consist of human or non-human cells, e.g., rodent cells, e.g., mouse or rat cells, rabbit cells, or pig cells. In preferred embodiments, the cell or cells include a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 transgene, e.g., a heterologous form of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751, e.g., a gene derived from humans (in the case of a non-human cell). The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 transgene can be misexpressed, e.g., overexpressed or underexpressed. In other preferred embodiments, the cell or cells include a gene which misexpresses an endogenous 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751, e.g., a gene the expression of which is disrupted, e.g., a knockout. Such cells can serve as a model for studying disorders which are related to mutated or misexpressed 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 alleles or for use in drug screening.

[0622] In another aspect, the invention features, a human cell, e.g., a hematopoietic stem cell, transformed with nucleic acid which encodes a subject 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide.

[0623] Also provided are cells, preferably human cells, e.g., human hematopoietic or fibroblast cells, in which an endogenous 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 is under the control of a regulatory sequence that does not normally control the expression of the endogenous 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene. The expression characteristics of an endogenous gene within a cell, e.g., a cell line or microorganism, can be modified by inserting a heterologous DNA regulatory element into the genome of the cell such that the inserted regulatory element is operably linked to the endogenous 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene. For example, an endogenous 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene which is "transcriptionally silent," e.g., not normally expressed, or expressed only at very low levels, can be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell. Techniques such as targeted homologous recombinations, can be used to insert the heterologous DNA as described in, e.g., Chappel, U.S. Pat. No. 5,272,071; WO 91/06667, published in May 16, 1991.

Transgenic Animals

[0624] The invention provides non-human transgenic animals. Such animals are useful for studying the function and/or activity of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein and for identifying and/or evaluating modulators of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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 or a rearrangement, e.g., a deletion of endogenous chromosomal DNA, which preferably is integrated into or occurs in the genome of the cells of a transgenic animal. A transgene can direct the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal, other transgenes, e.g., a knockout, reduce expression. Thus, a transgenic animal can be one in which an endogenous 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene has been altered by, e.g., 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.

[0625] 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 transgene of the invention to direct expression of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein to particular cells. A transgenic founder animal can be identified based upon the presence of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 transgene in its genome and/or expression of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein can further be bred to other transgenic animals carrying other transgenes.

[0626] 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins or polypeptides can be expressed in transgenic animals or plants, e.g., a nucleic acid encoding the protein or polypeptide can be introduced into the genome of an animal. In preferred embodiments the nucleic acid is placed under the control of a tissue specific promoter, e.g., a milk or egg specific promoter, and recovered from the milk or eggs produced by the animal. Suitable animals are mice, pigs, cows, goats, and sheep.

[0627] The invention also includes a population of cells from a transgenic animal, as discussed, e.g., below.

Uses

[0628] The nucleic acid molecules, proteins, protein homologs, and antibodies described herein can be used in one or more of the following methods: a) screening assays;

[0629] b) predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics); and c) methods of treatment (e.g., therapeutic and prophylactic).

[0630] The isolated nucleic acid molecules of the invention can be used, for example, to express a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein (e.g., via a recombinant expression vector in a host cell in gene therapy applications), to detect a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA (e.g., in a biological sample) or a genetic alteration in a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene, and to modulate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity, as described further below. The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins can be used to treat disorders characterized by insufficient or excessive production of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 substrate or production of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 inhibitors. In addition, the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins can be used to screen for naturally occurring 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 substrates, to screen for drugs or compounds which modulate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity, as well as to treat disorders characterized by insufficient or excessive production of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or production of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein forms which have decreased, aberrant or unwanted activity compared to 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 wild type protein (e.g., aberrant or deficient calcium channel activity, calcium/sodium antiporter activity, potassium channel activity, organic ion transporter activity, choline transporter activity, or other activity). Moreover, the anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibodies of the invention can be used to detect and isolate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins, regulate the bioavailability of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins, and modulate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity.

[0631] A method of evaluating a compound for the ability to interact with, e.g., bind, a subject 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide is provided. The method includes: contacting the compound with the subject 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide; and evaluating ability of the compound to interact with, e.g., to bind or form a complex with the subject 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide. This method can be performed in vitro, e.g., in a cell free system, or in vivo, e.g., in a two-hybrid interaction trap assay. This method can be used to identify naturally occurring molecules which interact with subject 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301,

58324, 55063, 52991, 59914, 59921 or 33751 polypeptide. It can also be used to find natural or synthetic inhibitors of subject 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide. Screening methods are discussed in more detail below.

Screening Assays:

[0632] The invention provides methods (also referred to herein as "screening assays") for identifying modulators, i.e., candidate or test compounds or agents (e.g., proteins, peptides, peptidomimetics, peptoids, small molecules or other drugs) which bind to 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins, have a stimulatory or inhibitory effect on, for example, 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 substrate. Compounds thus identified can be used to modulate the activity of target gene products (e.g., 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 genes) in a therapeutic protocol, to elaborate the biological function of the target gene product, or to identify compounds that disrupt normal target gene interactions.

[0633] In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or polypeptide or a 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or polypeptide or a biologically active portion thereof.

[0634] 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; peptoid libraries (libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone which are resistant to enzymatic degradation but which nevertheless remain bioactive; see, e.g., Zuckermann et al. (1994) *J. Med. Chem.* 37:2678-85); 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 and peptoid library approaches are limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des.* 12:145).

[0635] 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.* 90:6909-13; Erb et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:11422-426; Zuckermann et al. (1994). *J. Med. Chem.* 37:2678-85; 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.* 37:1233-51.

[0636] Libraries of compounds can 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, U.S. Pat. No. 5,223,409), spores (Ladner U.S. Pat. No. '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.).

[0637] In one embodiment, an assay is a cell-based assay in which a cell which expresses a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or biologically active portion thereof is contacted with a test compound, and the ability of the test compound to modulate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity is determined. Determining the ability of the test compound to modulate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity can be accomplished by monitoring, for example, calcium channel activity, calcium/sodium antiporter activity, potassium channel activity, organic ion transporter activity, choline transporter activity, or other activity. The cell, for example, can be of mammalian origin, e.g., human.

[0638] The ability of the test compound to modulate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 binding to a compound, e.g., a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 substrate, or to bind to 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 can also be evaluated. This can be accomplished, for example, by coupling the compound, e.g., the substrate, with a radioisotope or enzymatic label such that binding of the compound, e.g., the substrate, to 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 can be determined by detecting the labeled compound, e.g., substrate, in a complex. Alternatively, 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 could be coupled with a radioisotope or enzymatic label to monitor the ability of a test compound to modulate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 binding to a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 substrate in a complex. For example, compounds (e.g., 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 substrates) can be labeled with ^{125}I, ^{14}C, ^{35}S or ^{3}H., either directly or indirectly, and the radioisotope detected by direct counting of radioemmission or by scintillation counting. Alternatively, compounds can 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.

[0639] The ability of a compound (e.g., a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 substrate) to interact

with 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 with or without the labeling of any of the interactants can be evaluated. For example, a microphysiometer can be used to detect the interaction of a compound with 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 without the labeling of either the compound or the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751. McConnell 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751.

[0640] In yet another embodiment, a cell-free assay is provided in which a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or biologically active portion thereof is evaluated. Preferred biologically active portions of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins to be used in assays of the present invention include fragments which participate in interactions with non-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecules, e.g., fragments with high surface probability scores.

[0641] Soluble and/or membrane-bound forms of isolated proteins (e.g., 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins or biologically active portions thereof) can be used in the cell-free assays of the invention. When membrane-bound forms of the protein are used, it may be desirable to utilize a solubilizing agent. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)$_n$, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

[0642] Cell-free assays involve preparing a reaction mixture of the target gene protein and the test compound under conditions and for a time sufficient to allow the two components to interact and bind, thus forming a complex that can be removed and/or detected.

[0643] The interaction between two molecules can also be detected, e.g., using fluorescence energy transfer (FET) (see, for example, Lakowicz et al., U.S. Pat. No. 5,631,169; Stavrianopoulos, et al., U.S. Pat. No. 4,868,103). A fluorophore label on the first, 'donor' molecule is selected such that its emitted fluorescent energy will be absorbed by a fluorescent label on a second, 'acceptor' molecule, which in turn is able to fluoresce due to the absorbed energy. Alternately, the 'donor' protein molecule can simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the 'acceptor' molecule label can be differentiated from that of the 'donor'. Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, the spatial relationship between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the 'acceptor' molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means well known in the art (e.g., using a fluorimeter).

[0644] In another embodiment, determining the ability of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein to bind to a target molecule can be accomplished using real-time Biomolecular Interaction Analysis (BIA) (see, e.g., Sjolander and Urbaniczky (1991) *Anal. Chem.* 63:2338-2345 and Szabo et al. (1995) *Curr. Opin. Struct. Biol.* 5:699-705). "Surface plasmon resonance" or "BIA" detects biospecific interactions in real time, without labeling any of the interactants (e.g., BIAcore). Changes in the mass at the binding surface (indicative of a binding event) result in alterations of the refractive index of light near the surface (the optical phenomenon of surface plasmon resonance (SPR)), resulting in a detectable signal which can be used as an indication of real-time reactions between biological molecules.

[0645] In one embodiment, the target gene product or the test substance is anchored onto a solid phase. The target gene product/test compound complexes anchored on the solid phase can be detected at the end of the reaction. Preferably, the target gene product can be anchored onto a solid surface, and the test compound, (which is not anchored), can be labeled, either directly or indirectly, with detectable labels discussed herein.

[0646] It may be desirable to immobilize either 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751, an anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibody 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, or interaction of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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 microtiter 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/18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH).

[0647] Following incubation, the beads or microtiter 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 binding or activity determined using standard techniques.

[0648] Other techniques for immobilizing either a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or a target molecule on matrices include using conjugation of biotin and streptavidin. Biotinylated 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical).

[0649] In order to conduct the assay, the non-immobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (e.g., by washing) under conditions such that any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the previously non-immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously non-immobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific or selective for the immobilized component (the antibody, in turn, can be directly labeled or indirectly labeled with, e.g., a labeled anti-Ig antibody).

[0650] In one embodiment, this assay is performed utilizing antibodies reactive with 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or target molecules but which do not interfere with binding of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein to its target molecule. Such antibodies can be derivatized to the wells of the plate, and unbound target or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or target molecule.

[0651] Alternatively, cell free assays can be conducted in a liquid phase. In such an assay, the reaction products are separated from unreacted components, by any of a number of standard techniques, including but not limited to: differential centrifugation (see, for example, Rivas and Minton (1993) *Trends Biochem Sci* 18:284-7); chromatography (gel filtration chromatography, ion-exchange chromatography); electrophoresis (see, e.g., Ausubel et al., eds. (1999) *Current Protocols in Molecular Biology*, J. Wiley, New York); and immunoprecipitation (see, for example, Ausubel et al., eds. (1999) *Current Protocols in Molecular Biology*, J. Wiley, New York). Such resins and chromatographic techniques are known to one skilled in the art (see, e.g., Heegaard (1998) *J Mol Recognit* 11:141-8; Hage and Tweed (1997) *J Chromatogr B Biomed Sci Appl.* 699:499-525). Further, fluorescence energy transfer can also be conveniently utilized, as described herein, to detect binding without further purification of the complex from solution.

[0652] In a preferred embodiment, the assay includes contacting the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or biologically active portion thereof with a known compound which binds 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, wherein determining the ability of the test compound to interact with a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein includes determining the ability of the test compound to preferentially bind to 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 or biologically active portion thereof, or to modulate the activity of a target molecule, as compared to the known compound.

[0653] The target gene products of the invention can, in vivo, interact with one or more cellular or extracellular macromolecules, such as proteins. For the purposes of this discussion, such cellular and extracellular macromolecules are referred to herein as "binding partners." Compounds that disrupt such interactions can be useful in regulating the activity of the target gene product. Such compounds can include, but are not limited to molecules such as antibodies, peptides, and small molecules. The preferred target genes/products for use in this embodiment are the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 genes herein identified. In an alternative embodiment, the invention provides methods for determining the ability of the test compound to modulate the activity of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein through modulation of the activity of a downstream effector of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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.

[0654] To identify compounds that interfere with the interaction between the target gene product and its cellular or extracellular binding partner(s), a reaction mixture containing the target gene product and the binding partner is prepared, under conditions and for a time sufficient, to allow the two products to form complex. In order to test an inhibitory agent, the reaction mixture is provided in the presence and absence of the test compound. The test compound can be initially included in the reaction mixture, or can be added at a time subsequent to the addition of the target gene and its

cellular or extracellular binding partner. Control reaction mixtures are incubated without the test compound or with a placebo. The formation of any complexes between the target gene product and the cellular or extracellular binding partner is then detected. The formation of a complex in the control reaction, but not in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of the target gene product and the interactive binding partner.

[0655] Additionally, complex formation within reaction mixtures containing the test compound and normal target gene product can also be compared to complex formation within reaction mixtures containing the test compound and mutant target gene product. This comparison can be important in those cases wherein it is desirable to identify compounds that disrupt interactions of mutant but not normal target gene products.

[0656] These assays can be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring either the target gene product or the binding partner onto a solid phase, and detecting complexes anchored on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction between the target gene products and the binding partners, e.g., by competition, can be identified by conducting the reaction in the presence of the test substance. Alternatively, test compounds that disrupt preformed complexes, e.g., compounds with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed. The various formats are briefly described below.

[0657] In a heterogeneous assay system, either the target gene product or the interactive cellular or extracellular binding partner, is anchored onto a solid surface (e.g., a microtiter plate), while the non-anchored species is labeled, either directly or indirectly. The anchored species can be immobilized by non-covalent or covalent attachments. Alternatively, an immobilized antibody specific or selective for the species to be anchored can be used to anchor the species to the solid surface.

[0658] In order to conduct the assay, the partner of the immobilized species is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted components are removed (e.g., by washing) and any complexes formed will remain immobilized on the solid surface. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the non-immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific or selective for the initially non-immobilized species (the antibody, in turn, can be directly labeled or indirectly labeled with, e.g., a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds that inhibit complex formation or that disrupt preformed complexes can be detected.

[0659] Alternatively, the reaction can be conducted in a liquid phase in the presence or absence of the test compound, the reaction products separated from unreacted components, and complexes detected; e.g., using an immobilized antibody specific or selective for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific or selective for the other partner to detect anchored complexes. Again, depending upon the order of addition of reactants to the liquid phase, test compounds that inhibit complex or that disrupt preformed complexes can be identified.

[0660] In an alternate embodiment of the invention, a homogeneous assay can be used. For example, a preformed complex of the target gene product and the interactive cellular or extracellular binding partner product is prepared in that either the target gene products or their binding partners are labeled, but the signal generated by the label is quenched due to complex formation (see, e.g., U.S. Pat. No. 4,109,496 that utilizes this approach for immunoassays). The addition of a test substance that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances that disrupt target gene product-binding partner interaction can be identified.

[0661] In yet another aspect, the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Pat. 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 ("18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-binding proteins" or "18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-bp") and are involved in 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity. Such 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-bps can be activators or inhibitors of signals by the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 proteins or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 targets as, for example, downstream elements of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-mediated signaling pathway.

[0662] 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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. (Alternatively the: 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein can be the fused to the activator domain.) If the "bait" and the "prey" proteins are able to interact, in vivo, forming a 18607,

15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein.

[0663] In another embodiment, modulators of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression are identified. For example, a cell or cell free mixture is contacted with a candidate compound and the expression of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA or protein evaluated relative to the level of expression of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA or protein in the absence of the candidate compound. When expression of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA or protein is greater in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA or protein expression. Alternatively, when expression of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA or protein expression. The level of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA or protein expression can be determined by methods described herein for detecting 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA or protein.

[0664] 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein can be confirmed in vivo, e.g., in an animal such as an animal model for aberrant or deficient calcium channel activity, calcium/sodium antiporter activity, potassium channel activity, organic ion transporter activity, or choline transporter activity.

[0665] 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 (e.g., a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 modulating agent, an antisense 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid molecule, a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-specific antibody, or a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-binding partner) in an appropriate animal model to determine the efficacy, toxicity, side effects, or mechanism of action, of treatment with such an agent. Furthermore, novel agents identified by the above-described screening assays can be used for treatments as described herein.

Detection Assays

[0666] Portions or fragments of the nucleic acid sequences identified herein can be used as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome e.g., to locate gene regions associated with genetic disease or to associate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 with a disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

Chromosome Mapping

[0667] The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleotide sequences or portions thereof can be used to map the location of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 genes on a chromosome. This process is called chromosome mapping. Chromosome mapping is useful in correlating the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequences with genes associated with disease.

[0668] Briefly, 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleotide sequences. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequences will yield an amplified fragment.

[0669] A panel of somatic cell hybrids in which each cell line contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, can allow easy mapping of individual genes to specific human chromosomes. (D'Eustachio et al. (1983) *Science* 220:919-924).

[0670] Other mapping strategies e.g., in situ hybridization (described in Fan et al. (1990) *Proc. Natl. Acad. Sci. USA,* 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries can be used to map 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 to a chromosomal location.

[0671] Fluorescence in situ hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases will suffice to get good results at a reasonable amount of time. For a review of this technique, see Verma et al. (1988) Human

Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York).

[0672] Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

[0673] Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in McKusick, *Mendelian Inheritance in Man*, available on-line through Johns Hopkins University Welch Medical Library). The relationship between a gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland et al. (1987) *Nature,* 325:783-787.

[0674] Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

Tissue Typing

[0675] 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequences can be used to identify individuals from biological samples using, e.g., restriction fragment length polymorphism (RFLP). In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, the fragments separated, e.g., in a Southern blot, and probed to yield bands for identification. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Pat. No. 5,272,057).

[0676] Furthermore, the sequences of the present invention can also be used to determine the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleotide sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it. Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences.

[0677] Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1, 5, 13, 18, 21, 24, 27, 32, 35, 38, 53, 56, 60, 69, 72 or 81 can provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in 0 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

[0678] If a panel of reagents from 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleotide sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

Use of Partial 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequences in Forensic Biology

[0679] DNA-based identification techniques can also be used in forensic biology. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

[0680] The sequences of the present invention can be used to provide polynucleotide reagents, e.g., PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another, "identification marker" (i.e. another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NO:1, 5, 13, 18, 21, 24, 27, 32, 35, 38, 53, 56, 60, 69, 72 or 81 (e.g., fragments derived from the noncoding regions of SEQ ID NO:1, 5, 13, 18, 21, 24, 27, 32, 35, 38, 53, 56, 60, 69, 72 or 81 having a length of at least 20 bases, preferably at least 30 bases) are particularly appropriate for this use.

[0681] The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleotide sequences described herein can further be

used to provide polynucleotide reagents, e.g., labeled or labelable probes which can be used in, for example, an in situ hybridization technique, to identify a specific tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 probes can be used to identify tissue by species and/or by organ type.

[0682] In a similar fashion, these reagents, e.g., 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 primers or probes can be used to screen tissue culture for contamination (i.e. screen for the presence of a mixture of different types of cells in a culture).

Predictive Medicine

[0683] 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.

[0684] Generally, the invention provides, a method of determining if a subject is at risk for a disorder related to a lesion in or the misexpression of a gene which encodes 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751.

[0685] Such disorders include, e.g., a disorder associated with the misexpression of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene; a cellular proliferation or differentiation disorder, a cardiovascular, endothelial, breast, lung, colon, prostate, pancreas, brain, blood vessel, platelet, bone, immune, metabolic, kidney, ovarian, viral, pain, liver, skeletal muscle testicular, eye, hormonal, neurological, neurodegenerative, or angiogenic disorder.

[0686] The method includes one or more of the following: detecting, in a tissue of the subject, the presence or absence of a mutation which affects the expression of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene, or detecting the presence or absence of a mutation in a region which controls the expression of the gene, e.g., a mutation in the 5' control region; detecting, in a tissue of the subject, the presence or absence of a mutation which alters the structure of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene; detecting, in a tissue of the subject, the misexpression of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene, at the mRNA level, e.g., detecting a non-wild type level of an mRNA; or detecting, in a tissue of the subject, the misexpression of the gene, at the protein level, e.g., detecting a non-wild type level of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide.

[0687] In preferred embodiments the method includes: ascertaining the existence of at least one of: a deletion of one or more nucleotides from the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene; an insertion of one or more nucleotides into the gene, a point mutation, e.g., a substitution of one or more nucleotides of the gene, a gross chromosomal rearrangement of the gene, e.g., a translocation, inversion, or deletion.

[0688] For example, detecting the genetic lesion can include: (i) providing a probe/primer including an oligonucleotide containing a region of nucleotide sequence which hybridizes to a sense or antisense sequence from SEQ ID NO:1, 5, 13, 18, 21, 24, 27, 32, 35, 38, 53, 56, 60, 69, 72 or 81, or naturally occurring mutants thereof or 5' or 3' flanking sequences naturally associated with the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene; (ii) exposing the probe/primer to nucleic acid of the tissue; and detecting, by hybridization, e.g., in situ hybridization, of the probe/primer to the nucleic acid, the presence or absence of the genetic lesion.

[0689] In preferred embodiments detecting the misexpression includes ascertaining the existence of at least one of: an alteration in the level of a messenger RNA transcript of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene; the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene; or a non-wild type level of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751.

[0690] Methods of the invention can be used prenatally or to determine if a subject's offspring will be at risk for a disorder.

[0691] In preferred embodiments the method includes determining the structure of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene, an abnormal structure being indicative of risk for the disorder.

[0692] In preferred embodiments the method includes contacting a sample from the subject with an antibody to the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or a nucleic acid, which hybridizes specifically with the gene. These and other embodiments are discussed below.

Diagnostic and Prognostic Assays

[0693] The presence, level, or absence of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or nucleic acid in a biological sample can be evaluated by obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or nucleic acid (e.g., mRNA, genomic DNA) that encodes 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein such that the presence of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or nucleic acid is detected in the biological sample. The term "biological sample" includes tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. A preferred biological sample is serum. The level of expression of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene can be measured in a number of ways, including, but not limited to: measuring the mRNA encoded by the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 genes; measuring the amount of protein encoded by the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324,

55063, 52991, 59914, 59921 or 33751 genes; or measuring the activity of the protein encoded by the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 genes.

[0694] The level of mRNA corresponding to the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene in a cell can be determined both by in situ and by in vitro formats.

[0695] The isolated mRNA can be used in hybridization or amplification assays that include, but are not limited to, Southern or Northern analyses, polymerase chain reaction analyses and probe arrays. One preferred diagnostic method for the detection of mRNA levels involves contacting the isolated mRNA with a nucleic acid molecule (probe) that can hybridize to the mRNA encoded by the gene being detected. The nucleic acid probe can be, for example, a full-length 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid, such as the nucleic acid of SEQ ID NO:1, 5, 13, 18, 21, 24, 27, 32, 35, 38, 53, 56, 60, 69, 72 or 81, or a portion thereof, such as an oligonucleotide of at least 7, 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays are described herein.

[0696] In one format, mRNA (or cDNA) is immobilized on a surface and contacted with the probes, for example by running the isolated mRNA on an agarose gel and transferring the mRNA from the gel to a membrane, such as nitrocellulose. In an alternative format, the probes are immobilized on a surface and the mRNA (or cDNA) is contacted with the probes, for example, in a two-dimensional gene chip array. A skilled artisan can adapt known mRNA detection methods for use in detecting the level of mRNA encoded by the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 genes.

[0697] The level of mRNA in a sample that is encoded by one of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 can be evaluated with nucleic acid amplification, e.g., by rtPCR (Mullis (1987) U.S. Pat. No. 4,683,202), ligase chain reaction (Barany (1991) *Proc. Natl. Acad. Sci. USA* 88:189-193), self sustained sequence replication (Guatelli et al., (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh et al., (1989), *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi et al., (1988) *Bio/Technology* 6:1197), rolling circle replication (Lizardi et al., U.S. Pat. No. 5,854,033) or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques known in the art. As used herein, amplification primers are defined as being a pair of nucleic acid molecules that can anneal to 5' or 3' regions of a gene (plus and minus strands, respectively, or vice-versa) and contain a short region in between. In general, amplification primers are from about 10 to 30 nucleotides in length and flank a region from about 50 to 200 nucleotides in length. Under appropriate conditions and with appropriate reagents, such primers permit the amplification of a nucleic acid molecule comprising the nucleotide sequence flanked by the primers.

[0698] For in situ methods, a cell or tissue sample can be prepared/processed and immobilized on a support, typically a glass slide, and then contacted with a probe that can hybridize to mRNA that encodes the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene being analyzed.

[0699] In another embodiment, the methods further contacting a control sample with a compound or agent capable of detecting 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA, or genomic DNA, and comparing the presence of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA or genomic DNA in the control sample with the presence of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA or genomic DNA in the test sample.

[0700] A variety of methods can be used to determine the level of protein encoded by 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751. In general, these methods include contacting an agent that selectively binds to the protein, such as an antibody with a sample, to evaluate the level of protein in the sample. In a preferred embodiment, the antibody bears 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 a detectable substance. Examples of detectable substances are provided herein.

[0701] The detection methods can be used to detect 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein in a biological sample in vitro as well as in vivo. In vitro techniques for detection of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein include enzyme linked immunosorbent assays (ELISAs), immunoprecipitations, immunofluorescence, enzyme immunoassay (EIA), radioimmunoassay (RIA), and Western blot analysis. In vivo techniques for detection of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein include introducing into a subject a labeled anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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.

[0702] In another embodiment, the methods further include contacting the control sample with a compound or agent capable of detecting 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein, and comparing the presence of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein in the control sample with the presence of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein in the test sample.

[0703] The invention also includes kits for detecting the presence of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 in a biological sample. For example, the kit can include

a compound or agent capable of detecting 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or mRNA in a biological sample; and a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or nucleic acid.

[0704] For antibody-based kits, the kit can include: (1) a first antibody (e.g., attached to a solid support) which binds to a polypeptide corresponding to a marker of the invention; and, optionally, (2) a second, different antibody which binds to either the polypeptide or the first antibody and is conjugated to a detectable agent.

[0705] For oligonucleotide-based kits, the kit can include: (1) an oligonucleotide, e.g., a detectably labeled oligonucleotide, which hybridizes to a nucleic acid sequence encoding a polypeptide corresponding to a marker of the invention or (2) a pair of primers useful for amplifying a nucleic acid molecule corresponding to a marker of the invention. The kit can also includes a buffering agent, a preservative, or a protein stabilizing agent. The kit can also includes components necessary for detecting the detectable agent (e.g., an enzyme or a substrate). The kit can also contain a control sample or a series of control samples which can be assayed and compared to the test sample contained. Each component of the kit can be enclosed within an individual container and all of the various containers can be within a single package, along with instructions for interpreting the results of the assays performed using the kit.

[0706] The diagnostic methods described herein can identify subjects having, or at risk of developing, a disease or disorder associated with misexpressed or aberrant or unwanted 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression or activity. As used herein, the term "unwanted" includes an unwanted phenomenon involved in a biological response such as pain or deregulated cell proliferation.

[0707] In one embodiment, a disease or disorder associated with aberrant or unwanted 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression or activity is identified. A test sample is obtained from a subject and 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or nucleic acid (e.g., mRNA or genomic DNA) is evaluated, wherein the level, e.g., the presence or absence, of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant or unwanted 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest, including a biological fluid (e.g., serum), cell sample, or tissue.

[0708] The prognostic assays described herein can be used to determine whether a subject can be administered an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant or unwanted 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a cellular proliferation or differentiation disorder, a cardiovascular, endothelial, breast, lung, colon, prostate, pancreas, brain, blood vessel, platelet, bone, immune, metabolic, kidney, ovarian, viral, pain, liver, skeletal muscle testicular, eye, hormonal, neurological, neurodegenerative or angiogenic disorder.

[0709] The methods of the invention can also be used to detect genetic alterations in a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene, thereby determining if a subject with the altered gene is at risk for a disorder characterized by misregulation in 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein activity or nucleic acid expression, such as a a cellular proliferation or differentiation disorder, a cardiovascular, endothelial, breast, lung, colon, prostate, pancreas, brain, blood vessel, platelet, bone, immune, metabolic, kidney, ovarian, viral, pain, liver, skeletal muscle testicular, eye, hormonal neurological, neurodegenerative or angiogenic disorder. In preferred embodiments, the methods include detecting, in a sample from the subject, the presence or absence of a genetic alteration characterized by at least one of an alteration affecting the integrity of a gene encoding a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-protein, or the mis-expression of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene. 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene; 2) an addition of one or more nucleotides to a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene; 3) a substitution of one or more nucleotides of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene, 4) a chromosomal rearrangement of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene; 5) an alteration in the level of a messenger RNA transcript of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene, 6) aberrant modification of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene, 8) a non-wild type level of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-protein, 9) allelic loss of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene, and 10) inappropriate post-translational modification of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-protein.

[0710] An alteration can be detected without a probe/primer in a polymerase chain reaction, such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR), the latter of which can be particularly useful for detecting point mutations in the 18607, 15603, 69318, 12303,

48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-gene. This method can include the steps of collecting a sample of cells from a subject, isolating nucleic acid (e.g., genomic, mRNA or both) from the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene under conditions such that hybridization and amplification of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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. Alternatively, other amplification methods described herein or known in the art can be used.

[0711] In another embodiment, mutations in a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene from a sample cell can be identified by detecting 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, e.g., 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. Pat. No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

[0712] In other embodiments, genetic mutations in 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 can be identified by hybridizing a sample and control nucleic acids, e.g., DNA or RNA, two dimensional arrays, e.g., chip based arrays. Such arrays include a plurality of addresses, each of which is positionally distinguishable from the other. A different probe is located at each address of the plurality. The arrays can have a high density of addresses, e.g., can contain hundreds or thousands of oligonucleotides probes (Cronin et al. (1996) *Human Mutation* 7: 244-255; Kozal et al. (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, M. T. et al. 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 the identification of point mutations. This step is followed by a second hybridization array that allows 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.

[0713] In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene and detect mutations by comparing the sequence of the sample 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 with the corresponding wild-type (control) sequence. Automated sequencing procedures can be utilized when performing the diagnostic assays (Naeve et al. (1995) *Biotechniques* 19:448-53), including sequencing by mass spectrometry.

[0714] Other methods for detecting mutations in the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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; Cotton et al. (1988) *Proc. Natl. Acad Sci USA* 85:4397; Saleeba et al. (1992) *Methods Enzymol.* 217:286-295).

[0715] 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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; U.S. Pat. No. 5,459,039).

[0716] In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 genes. For example, single strand conformation polymorphism (SSCP) can 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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 can be labeled or detected with labeled probes. The sensitivity of the assay can 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).

[0717] 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 insure 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).

[0718] Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension (Saiki et al. (1986) *Nature* 324:163); Saiki et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:6230).

[0719] Alternatively, allele specific amplification technology which depends on selective PCR amplification can be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification can 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 can also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci. USA* 88:189-93). 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.

[0720] The methods described herein can be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which can be conveniently used, e.g., in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene.

Use of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 Molecules as Surrogate Markers

[0721] The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecules of the invention are also useful as markers of disorders or disease states, as markers for precursors of disease states, as markers for predisposition of disease states, as markers of drug activity, or as markers of the pharmacogenomic profile of a subject. Using the methods described herein, the presence, absence and/or quantity of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecules of the invention can be detected, and can be correlated with one or more biological states in vivo. For example, the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecules of the invention can serve as surrogate markers for one or more disorders or disease states or for conditions leading up to disease states. As used herein, a "surrogate marker" is an objective biochemical marker which correlates with the absence or presence of a disease or disorder, or with the progression of a disease or disorder (e.g., with the presence or absence of a tumor). The presence or quantity of such markers is independent of the disease. Therefore, these markers can serve to indicate whether a particular course of treatment is effective in lessening a disease state or disorder. Surrogate markers are of particular use when the presence or extent of a disease state or disorder is difficult to assess through standard methodologies (e.g., early stage tumors), or when an assessment of disease progression is desired before a potentially dangerous clinical endpoint is reached (e.g., an assessment of cardiovascular disease can be made using cholesterol levels as a surrogate marker, and an analysis of HIV infection can be made using HIV RNA levels as a surrogate marker, well in advance of the undesirable clinical outcomes of myocardial infarction or fully-developed AIDS). Examples of the use of surrogate markers in the art include: Koomen et al. (2000) *J. Mass. Spectrom.* 35: 258-264; and James (1994) *AIDS Treatment News Archive* 209.

[0722] The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecules of the invention are also useful as pharmacodynamic markers. As used herein, a "pharmacodynamic marker" is an objective biochemical marker which correlates specifically with drug effects. The presence or quantity of a pharmacodynamic marker is not related to the disease state or disorder for which the drug is being administered; therefore, the presence or quantity of the marker is indicative of the presence or activity of the drug in a subject. For example, a pharmacodynamic marker can be indicative of the concentration of the drug in a biological tissue, in that the marker is either expressed or transcribed or not expressed or transcribed in that tissue in relationship to the level of the drug. In this fashion, the distribution or uptake of the drug can be monitored by the pharmacodynamic marker. Similarly, the presence or quantity of the pharmacodynamic marker can be related to the presence or quantity of the metabolic product of a drug, such that the presence or quantity of the marker is indicative of the relative breakdown rate of the drug in vivo. Pharmacodynamic markers are of particular use in increasing the sensitivity of detection of drug effects, particularly when the drug is administered in low doses. Since even a small amount of a drug can be sufficient to activate multiple rounds of marker (e.g., a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 marker) transcription or expression, the amplified marker can be in a quantity which is more readily detectable than the drug itself. Also, the marker can be more easily detected due to the nature of the marker itself; for example, using the methods described herein, anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibodies can be employed in an immune-based detection system for a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein marker, or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-specific radiolabeled probes can be used to detect a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA marker. Furthermore, the use of a pharmacodynamic marker can offer mechanism-based prediction of risk due to drug treatment beyond the range of possible direct observations. Examples of the use of pharmacodynamic markers in the art include: Matsuda et al. U.S. Pat. No. 6,033,862; Hattis et al. (1991) *Env. Health Perspect.* 90: 229-238; Schentag (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S21-S24; and Nicolau (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S16-S20.

[0723] The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecules of the invention are also useful as pharmacogenomic markers. As used herein, a "pharmacogenomic marker" is an objective biochemical marker which correlates

with a specific clinical drug response or susceptibility in a subject (see, e.g., McLeod et al. (1999) *Eur. J. Cancer* 35:1650-1652). The presence or quantity of the pharmacogenomic marker is related to the predicted response of the subject to a specific drug or class of drugs prior to administration of the drug. By assessing the presence or quantity of one or more pharmacogenomic markers in a subject, a drug therapy which is most appropriate for the subject, or which is predicted to have a greater degree of success, can be selected. For example, based on the presence or quantity of RNA, or protein (e.g., 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or RNA) for specific tumor markers in a subject, a drug or course of treatment can be selected that is optimized for the treatment of the specific tumor likely to be present in the subject. Similarly, the presence or absence of a specific sequence mutation in 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 dNA can correlate with a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 drug response. The use of pharmacogenomic markers therefore permits the application of the most appropriate treatment for each subject without having to administer the therapy.

Pharmaceutical Compositions

[0724] The nucleic acid and polypeptides, fragments thereof, as well as anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions. Such compositions typically include the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" includes solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Supplementary active compounds can also be incorporated into the compositions.

[0725] A pharmaceutical composition 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.

[0726] 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, N.J.) 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 should 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, 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.

[0727] Sterile injectable solutions can be prepared by incorporating the active compound 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.

[0728] Oral compositions generally include an inert diluent or an edible carrier. 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, e.g., gelatin capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash. 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.

[0729] 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.

[0730] 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.

[0731] The compounds 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.

[0732] In one embodiment, the active compounds 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. Pat. No. 4,522,811.

[0733] It is 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.

[0734] Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds which exhibit high therapeutic indices are preferred. While compounds that exhibit toxic side effects can be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

[0735] 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 compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage can vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (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 can be measured, for example, by high performance liquid chromatography.

[0736] 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 protein or polypeptide can be administered 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. The skilled artisan will appreciate that certain factors can influence the dosage and timing 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, unconjugated or conjugated as described herein, can include a single treatment or, preferably, can include a series of treatments.

[0737] For antibodies, the preferred dosage is 0.1 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg). If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is usually appropriate. Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (e.g., into the brain). A method for lipidation of antibodies is described by Cruikshank et al. ((1997) *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193).

[0738] The present invention encompasses agents which modulate expression or activity. An agent can, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics (e.g., peptoids), 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.

[0739] 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. 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 can, 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.

[0740] The nucleic acid molecules 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. Pat. No. 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.

[0741] The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

Methods of Treatment:

[0742] The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant or unwanted 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression or activity. As used herein, the term "treatment" 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, a symptom of disease or a predisposition toward a disease, with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect the disease, the symptoms of disease or the predisposition toward disease. A therapeutic agent includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides.

[0743] With regards to both prophylactic and therapeutic methods of treatment, such treatments can 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 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 individual's prophylactic or therapeutic treatment with either the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecules of the present invention or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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.

[0744] In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant or unwanted 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression or activity, by administering to the subject a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 or an agent which modulates 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression or at least one 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity. Subjects at risk for a disease which is caused or contributed to by aberrant or unwanted 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 aberrance, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 aberrance, for example, a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751, 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 agonist or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

[0745] It is possible that some 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 disorders can be caused, at least in part, by an abnormal level of gene product, or by the presence of a gene product exhibiting abnormal activity. As such, the reduction in the level and/or activity of such gene products would bring about the amelioration of disorder symptoms.

[0746] The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecules can act as novel diagnostic targets and therapeutic agents for controlling one or more of a cellular proliferation and/or differentiation disorder, a cardiovascular, endothelial, breast, lung, colon, prostate, pancreas, brain, blood vessel, platelet, bone, immune, metabolic, kidney, ovarian, viral, pain, liver, skeletal muscle testicular, eye, hormonalneurological, neurodegenerative or angiogenic disorder, all of which are described above.

[0747] As discussed, successful treatment of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 disorders can be brought about by techniques that serve to inhibit the expression or activity of target gene products. For example, compounds, e.g., an agent identified using an assays described above, that proves to exhibit negative modulatory activity, can be used in accordance with the invention to prevent and/or ameliorate symptoms of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 disorders. Such molecules can include, but are not limited to peptides, phosphopeptides, small organic or inorganic molecules, or antibodies (including, for example, polyclonal, monoclonal, humanized, human, anti-idiotypic,

chimeric or single chain antibodies, and Fab, $F(ab')_2$ and Fab expression library fragments, scFV molecules, and epitope-binding fragments thereof).

[0748] Further, antisense and ribozyme molecules that inhibit expression of the target gene can also be used in accordance with the invention to reduce the level of target gene expression, thus effectively reducing the level of target gene activity. Still further, triple helix molecules can be utilized in reducing the level of target gene activity. Antisense, ribozyme and triple helix molecules are discussed above.

[0749] It is possible that the use of antisense, ribozyme, and/or triple helix molecules to reduce or inhibit mutant gene expression can also reduce or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by normal target gene alleles, such that the concentration of normal target gene product present can be lower than is necessary for a normal phenotype. In such cases, nucleic acid molecules that encode and express target gene polypeptides exhibiting normal target gene activity can be introduced into cells via gene therapy method. Alternatively, in instances in that the target gene encodes an extracellular protein, it can be preferable to co-administer normal target gene protein into the cell or tissue in order to maintain the requisite level of cellular or tissue target gene activity.

[0750] Another method by which nucleic acid molecules can be utilized in treating or preventing a disease characterized by 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression is through the use of aptamer molecules specific for 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein. Aptamers are nucleic acid molecules having a tertiary structure which permits them to specifically or selectively bind to protein ligands (see, e.g., Osborne et al. (1997) *Curr. Opin. Chem. Biol.* 1: 5-9; and Patel (1997) *Curr Opin Chem Biol* 1:32-46). Since nucleic acid molecules can in many cases be more conveniently introduced into target cells than therapeutic protein molecules can be, aptamers offer a method by which 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein activity can be specifically decreased without the introduction of drugs or other molecules which can have pluripotent effects.

[0751] Antibodies can be generated that are both specific for target gene product and that reduce target gene product activity. Such antibodies can, therefore, by administered in instances whereby negative modulatory techniques are appropriate for the treatment of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 disorders. For a description of antibodies, see the Antibody section above.

[0752] In circumstances wherein injection of an animal or a human subject with a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or epitope for stimulating antibody production is harmful to the subject, it is possible to generate an immune response against 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 through the use of anti-idiotypic antibodies (see, for example, Herlyn (1999) *Ann Med* 31:66-78; and Bhattacharya-Chatterjee and Foon (1998) *Cancer Treat Res.* 94:51-68). If an anti-idiotypic antibody is introduced into a mammal or human subject, it should stimulate the production of anti-anti-idiotypic antibodies, which should be specific to the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein.

[0753] Vaccines directed to a disease characterized by 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression can also be generated in this fashion.

[0754] In instances where the target antigen is intracellular and whole antibodies are used, internalizing antibodies can be preferred. Lipofectin or liposomes can be used to deliver the antibody or a fragment of the Fab region that binds to the target antigen into cells. Where fragments of the antibody are used, the smallest inhibitory fragment that binds to the target antigen is preferred. For example, peptides having an amino acid sequence corresponding to the Fv region of the antibody can be used. Alternatively, single chain neutralizing antibodies that bind to intracellular target antigens can also be administered. Such single chain antibodies can be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population (see e.g., Marasco et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:7889-7893).

[0755] The identified compounds that inhibit target gene expression, synthesis and/or activity can be administered to a patient at therapeutically effective doses to prevent, treat or ameliorate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 disorders. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms of the disorders. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures as described above.

[0756] 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 compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage can vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (i.e., the concentration of the test compound that 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 can be measured, for example, by high performance liquid chromatography.

[0757] Another example of determination of effective dose for an individual is the ability to directly assay levels of "free" and "bound" compound in the serum of the test subject. Such assays can utilize antibody mimics and/or "biosensors" that have been created through molecular imprinting techniques. The compound which is able to modulate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity is used as a template, or "imprinting molecule", to spatially organize polymerizable monomers prior to their polymerization with catalytic reagents. The subsequent removal of the imprinted molecule leaves a polymer matrix which contains a repeated "negative image" of the compound and is able to selectively rebind the molecule under biological assay conditions. A detailed review of this technique can be seen in Ansell et al

(1996) *Current Opinion in Biotechnology* 7:89-94 and in Shea (1994) *Trends in Polymer Science* 2:166-173. Such "imprinted" affinity matrixes are amenable to ligand-binding assays, whereby the immobilized monoclonal antibody component is replaced by an appropriately imprinted matrix. An example of the use of such matrixes in this way can be seen in Vlatakis et al (1993) *Nature* 361:645-647. Through the use of isotope-labeling, the "free" concentration of compound which modulates the expression or activity of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 can be readily monitored and used in calculations of IC_{50}.

[0758] Such "imprinted" affinity matrixes can also be designed to include fluorescent groups whose photon-emitting properties measurably change upon local and selective binding of target compound. These changes can be readily assayed in real time using appropriate fiberoptic devices, in turn allowing the dose in a test subject to be quickly optimized based on its individual IC_{50}. An rudimentary example of such a "biosensor" is discussed in Kriz et al (1995) *Analytical Chemistry* 67:2142-2144.

[0759] Another aspect of the invention pertains to methods of modulating 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell with a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 or agent that modulates one or more of the activities of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein activity associated with the cell. An agent that modulates 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring target molecule of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein (e.g., a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 substrate or receptor), a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibody, a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 agonist or antagonist, a peptidomimetic of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 agonist or antagonist, or other small molecule.

[0760] In one embodiment, the agent stimulates one or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activities. Examples of such stimulatory agents include active 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein and a nucleic acid molecule encoding 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751. In another embodiment, the agent inhibits one or more 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activities. Examples of such inhibitory agents include antisense 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid molecules, anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibodies, and 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 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., up regulates or down regulates) 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression or activity. In another embodiment, the method involves administering a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression or activity.

[0761] Stimulation of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity is desirable in situations in which 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 is abnormally downregulated and/or in which increased 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity is likely to have a beneficial effect. For example, stimulation of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity is desirable in situations in which a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 is downregulated and/or in which increased 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity is likely to have a beneficial effect. Likewise, inhibition of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity is desirable in situations in which 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 is abnormally upregulated and/or in which decreased 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity is likely to have a beneficial effect.

Pharmacogenomics

[0762] The 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecules of the present invention, as well as agents, or modulators which have a stimulatory or inhibitory effect on 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity (e.g., 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) 18607, 15603, 69318, 12303,

48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disorders (e.g., aberrant or deficient calcium channel activity, calcium/sodium antiporter activity, potassium channel activity, organic ion transporter activity, or choline transporter activity) associated with aberrant or unwanted 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity.

[0763] In conjunction with such treatment, pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) can 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 can consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecule or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 modulator as well as tailoring the dosage and/or therapeutic regimen of treatment with a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecule or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 modulator.

[0764] 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 et al. (1996) *Clin. Exp. Pharmacol. Physiol.* 23:983-985 and Linder et al. (1997) *Clin. Chem.* 43:254-266. 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 dehydrogenase 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.

[0765] 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 can occur once per every 1000 bases of DNA. A SNP can be involved in a disease process, however, the vast majority can 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 can be common among such genetically similar individuals.

[0766] 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's target is known (e.g., a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein 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.

[0767] 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 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecule or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 modulator of the present invention) can give an indication whether gene pathways related to toxicity have been turned on.

[0768] 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 an individual. 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 with a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecule or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 modulator, such as a modulator identified by one of the exemplary screening assays described herein.

[0769] The present invention further provides methods for identifying new agents, or combinations, that are based on identifying agents that modulate the activity of one or more of the gene products encoded by one or more of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 genes of the present invention, wherein these products can be associated with resistance of the cells to a therapeutic agent. Specifically, the activity of the proteins encoded by the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 genes of the present invention can be used as a basis for identifying agents for overcoming agent resistance. By blocking the activity of one or more of the resistance proteins, target cells, e.g., human cells, will become sensitive to treatment with an agent to which the unmodified target cells were resistant.

[0770] Monitoring the influence of agents (e.g., drugs) on the expression or activity of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein can be applied in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene expres-

sion, protein levels, or upregulate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity, can be monitored in clinical trials of subjects exhibiting decreased 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene expression, protein levels, or downregulated 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene expression, protein levels, or downregulate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity, can be monitored in clinical trials of subjects exhibiting increased 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene expression, protein levels, or upregulated 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity. In such clinical trials, the expression or activity of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene, and preferably, other genes that have been implicated in, for example, a ion channel-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disorder can be used as a "read out" or markers of the phenotype of a particular cell.

Other Embodiments

[0771] In another aspect, the invention features a method of analyzing a plurality of capture probes. The method is useful, e.g., to analyze gene expression. The method includes: providing a two dimensional array having a plurality of addresses, each address of the plurality being positionally distinguishable from each other address of the plurality, and each address of the plurality having a unique capture probe, e.g., a nucleic acid or peptide sequence, wherein the capture probes are from a cell or subject which expresses 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 or from a cell or subject in which a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mediated response has been elicited; contacting the array with a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid (preferably purified), a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide (preferably purified), or an anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibody, and thereby evaluating the plurality of capture probes. Binding, e.g., in the case of a nucleic acid, hybridization with a capture probe at an address of the plurality, is detected, e.g., by a signal generated from a label attached to the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid, polypeptide, or antibody.

[0772] The capture probes can be a set of nucleic acids from a selected sample, e.g., a sample of nucleic acids derived from a control or non-stimulated tissue or cell.

[0773] The method can include contacting the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid, polypeptide, or antibody with a first array having a plurality of capture probes and a second array having a different plurality of capture probes. The results of each hybridization can be compared, e.g., to analyze differences in expression between a first and second sample. The first plurality of capture probes can be from a control sample, e.g., a wild type, normal, or non-diseased, non-stimulated, sample, e.g., a biological fluid, tissue, or cell sample. The second plurality of capture probes can be from an experimental sample, e.g., a mutant type, at risk, disease-state or disorder-state, or stimulated, sample, e.g., a biological fluid, tissue, or cell sample.

[0774] The plurality of capture probes can be a plurality of nucleic acid probes each of which specifically hybridizes, with an allele of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751. Such methods can be used to diagnose a subject, e.g., to evaluate risk for a disease or disorder, to evaluate suitability of a selected treatment for a subject, to evaluate whether a subject has a disease or disorder.

[0775] The method can be used to detect SNPs, as described above.

[0776] In another aspect, the invention features, a method of analyzing 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751, e.g., analyzing structure, function, or relatedness to other nucleic acid or amino acid sequences. The method includes: providing a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid or amino acid sequence; comparing the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence with one or more preferably a plurality of sequences from a collection of sequences, e.g., a nucleic acid or protein sequence database; to thereby analyze 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751.

[0777] The method can include evaluating the sequence identity between a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence and a database sequence. The method can be performed by accessing the database at a second site, e.g., over the internet. Preferred databases include GenBank™ and SwissProt.

[0778] In another aspect, the invention features, a set of oligonucleotides, useful, e.g., for identifying SNP's, or identifying specific alleles of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751. The set includes a plurality of oligonucleotides, each of which has a different nucleotide at an interrogation position, e.g., an SNP or the site of a mutation. In a preferred embodiment, the oligonucleotides of the plurality identical in sequence with one another (except for differences in length). The oligonucleotides can be provided with differential labels, such that an oligonucleotide which hybridizes to one allele provides a signal that is distinguishable from an oligonucleotides which hybridizes to a second allele.

[0779] The sequences of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecules are provided in a variety of mediums to facilitate use thereof. A sequence can be provided as a manufacture, other than an isolated nucleic acid or amino acid molecule, which contains a 18607, 15603, 69318, 12303,

48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 molecule. Such a manufacture can provide a nucleotide or amino acid sequence, e.g., an open reading frame, in a form which allows examination of the manufacture using means not directly applicable to examining the nucleotide or amino acid sequences, or a subset thereof, as they exist in nature or in purified form.

[0780] A 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleotide or amino acid sequence can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium that can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as compact disc and CD-ROM; electrical storage media such as RAM, ROM, EPROM, EEPROM, and the like; and general hard disks and hybrids of these categories such as magnetic/optical storage media. The medium is adapted or configured for having thereon 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence information of the present invention.

[0781] As used herein, the term "electronic apparatus" is intended to include any suitable computing or processing apparatus of other device configured or adapted for storing data or information. Examples of electronic apparatus suitable for use with the present invention include stand-alone computing apparatus; networks, including a local area network (LAN), a wide area network (WAN) Internet, Intranet, and Extranet; electronic appliances such as personal digital assistants (PDAs), cellular phones, pagers, and the like; and local and distributed processing systems.

[0782] As used herein, "recorded" refers to a process for storing or encoding information on the electronic apparatus readable medium. Those skilled in the art can readily adopt any of the presently known methods for recording information on known media to generate manufactures comprising the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence information.

[0783] A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleotide or amino acid sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. The skilled artisan can readily adapt any number of data processor structuring formats (e.g., text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

[0784] By providing the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleotide or amino acid sequences of the invention in computer readable form, the skilled artisan can routinely access the sequence information for a variety of purposes. For example, one skilled in the art can use the nucleotide or amino acid sequences of the invention in computer readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. A search is used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

[0785] The present invention therefore provides a medium for holding instructions for performing a method for determining whether a subject has a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder or a pre-disposition to a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder, wherein the method comprises the steps of determining 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence information associated with the subject and based on the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence information, determining whether the subject has a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder and/or recommending a particular treatment for the disease, disorder, or pre-disease condition.

[0786] The present invention further provides in an electronic system and/or in a network, a method for determining whether a subject has a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder or a pre-disposition to a disease associated with 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751, wherein the method comprises the steps of determining 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence information associated with the subject, and based on the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence information, determining whether the subject has a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder or a pre-disposition to a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder, and/or recommending a particular treatment for the disease, disorder, or pre-disease condition. The method may further comprise the step of receiving

phenotypic information associated with the subject and/or acquiring from a network phenotypic information associated with the subject.

[0787] The present invention also provides in a network, a method for determining whether a subject has a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder or a pre-disposition to a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder, said method comprising the steps of receiving 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence information from the subject and/or information related thereto, receiving phenotypic information associated with the subject, acquiring information from the network corresponding to 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 and/or corresponding to a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder, and based on one or more of the phenotypic information, the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 information (e.g., sequence information and/or information related thereto), and the acquired information, determining whether the subject has a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder or a pre-disposition to a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder, or pre-disease condition.

[0788] The present invention also provides a business method for determining whether a subject has a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder or a pre-disposition to a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder, said method comprising the steps of receiving information related to 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 (e.g., sequence information and/or information related thereto), receiving phenotypic information associated with the subject, acquiring information from the network related to 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 and/or related to a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder, and based on one or more of the phenotypic information, the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 information, and the acquired information, determining whether the subject has a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder or a pre-disposition to a calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder, or pre-disease condition.

[0789] The invention also includes an array comprising a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence of the present invention. The array can be used to assay expression of one or more genes in the array. In one embodiment, the array can be used to assay gene expression in a tissue to ascertain tissue specificity of genes in the array. In this manner, up to about 7600 genes can be simultaneously assayed for expression, one of which can be 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751. This allows a profile to be developed showing a battery of genes specifically expressed in one or more tissues.

[0790] In addition to such qualitative information, the invention allows the quantitation of gene expression. Thus, not only tissue specificity, but also the level of expression of a battery of genes in the tissue if ascertainable. Thus, genes can be grouped on the basis of their tissue expression per se and level of expression in that tissue. This is useful, for example, in ascertaining the relationship of gene expression in that tissue. Thus, one tissue can be perturbed and the effect on gene expression in a second tissue can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. Such a determination is useful, for example, to know the effect of cell-cell interaction at the level of gene expression. If an agent is administered therapeutically to treat one cell type but has an undesirable effect on another cell type, the invention provides an assay to determine the molecular basis of the undesirable effect and thus provides the opportunity to co-administer a counteracting agent or otherwise treat the undesired effect. Similarly, even within a single cell type, undesirable biological effects can be determined at the molecular level. Thus, the effects of an agent on expression of other than the target gene can be ascertained and counteracted.

[0791] In another embodiment, the array can be used to monitor the time course of expression of one or more genes in the array. This can occur in various biological contexts, as disclosed herein, for example development of a calcium channel, calcium/sodium antiporter, potassium channel, organic

ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder, progression of calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder, and processes, such a cellular transformation associated with the calcium channel, calcium/sodium antiporter, potassium channel, organic ion transporter or choline transporter-associated or another 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751-associated disease or disorder.

[0792] The array is also useful for ascertaining the effect of the expression of a gene on the expression of other genes in the same cell or in different cells (e.g., acertaining the effect of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression on the expression of other genes). This provides, for example, for a selection of alternate molecular targets for therapeutic intervention if the ultimate or downstream target cannot be regulated.

[0793] The array is also useful for ascertaining differential expression patterns of one or more genes in normal and abnormal cells. This provides a battery of genes (e.g., including 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751) that could serve as a molecular target for diagnosis or therapeutic intervention.

[0794] As used herein, a "target sequence" can be any DNA or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. Typical sequence lengths of a target sequence are from about 10 to 100 amino acids or from about 30 to 300 nucleotide residues. However, it is well recognized that commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

[0795] Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium for analysis and comparison to other sequences. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software include, but are not limited to, MacPattern (EMBL), BLASTN and BLASTX (NCBI).

[0796] Thus, the invention features a method of making a computer readable record of a sequence of a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence which includes recording the sequence on a computer readable matrix. In a preferred embodiment the record includes one or more of the following: identification of an ORF; identification of a domain, region, or site; identification of the start of transcription; identification of the transcription terminator; the full length amino acid sequence of the protein, or a mature form thereof; the 5' end of the translated region.

[0797] In another aspect, the invention features a method of analyzing a sequence. The method includes: providing a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence, or record, in computer readable form; comparing a second sequence to the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence; thereby analyzing a sequence. Comparison can include comparing to sequences for sequence identity or determining if one sequence is included within the other, e.g., determining if the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 sequence includes a sequence being compared. In a preferred embodiment the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 or second sequence is stored on a first computer, e.g., at a first site and the comparison is performed, read, or recorded on a second computer, e.g., at a second site. E.g., the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 or second sequence can be stored in a public or proprietary database in one computer, and the results of the comparison performed, read, or recorded on a second computer. In a preferred embodiment the record includes one or more of the following: identification of an ORF; identification of a domain, region, or site; identification of the start of transcription; identification of the transcription terminator; the full length amino acid sequence of the protein, or a mature form thereof; the 5' end of the translated region.

EXEMPLIFICATION

Example 1

Tissue Distribution of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA

[0798] Northern blot hybridizations with various RNA samples can be performed under standard conditions and washed under stringent conditions, i.e., 0.2×SSC at 65° C. A DNA probe corresponding to all or a portion of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 cDNA (SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83) or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 cDNA can be used. The DNA was radioactively labeled with ^{32}P-dCTP using the Prime-It Kit (Stratagene, La Jolla, Calif.) according to the instructions of the supplier. Filters containing mRNA from mouse hematopoietic and endocrine tissues, and cancer cell lines (Clontech, Palo Alto, Calif.) can be probed in ExpressHyb hybridization solution (Clontech) and washed at high stringency according to manufacturer's recommendations.

Example 2

Recombinant Expression of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 in Bacterial Cells

[0799] In this example, 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 is expressed as a recombinant glutathione-S-transferase (GST) fusion polypeptide in *E. coli*

and the fusion polypeptide is isolated and characterized. Specifically, 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 is fused to GST and this fusion polypeptide is expressed in *E. coli*, e.g., strain PEB199. Expression of the GST-18607, -15603, -69318, -12303, -48000, -52920, -5433, -38554, -57301, -58324, -55063, -52991, -59914, -59921 or -33751 fusion protein in PEB199 is induced with IPTG. The recombinant fusion polypeptide is purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography on glutathione beads. Using polyacrylamide gel electrophoretic analysis of the polypeptide purified from the bacterial lysates, the molecular weight of the resultant fusion polypeptide is determined.

Example 3

Expression of Recombinant 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 Protein in COS Cells

[0800] To express the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene in COS cells, the pcDNA/Amp vector by Invitrogen Corporation (San Diego, Calif.) is used. This vector contains an SV40 origin of replication, an ampicillin resistance gene, an *E. coli* replication origin, a CMV promoter followed by a polylinker region, and an SV40 intron and polyadenylation site. A DNA fragment encoding the entire 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 protein and an HA tag (Wilson et al. (1984) *Cell* 37:767) or a FLAG tag fused in-frame to its 3' end of the fragment is cloned into the polylinker region of the vector, thereby placing the expression of the recombinant protein under the control of the CMV promoter.

[0801] To construct the plasmid, the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 dNA sequence is amplified by PCR using two primers. The 5' primer contains the restriction site of interest followed by approximately twenty nucleotides of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 coding sequence starting from the initiation codon; the 3' end sequence contains complementary sequences to the other restriction site of interest, a translation stop codon, the HA tag or FLAG tag and the last 20 nucleotides of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 coding sequence. The PCR amplified fragment and the pcDNA/Amp vector are digested with the appropriate restriction enzymes and the vector is dephosphorylated using the CIAP enzyme (New England Biolabs, Beverly, Mass.). Preferably the two restriction sites chosen are different so that the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene is inserted in the correct orientation. The ligation mixture is transformed into *E. coli* cells (strains HB101, DH5α, SURE, available from Stratagene Cloning Systems, La Jolla, Calif., can be used), the transformed culture is plated on ampicillin media plates, and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct fragment.

[0802] COS cells are subsequently transfected with the 18607-, 15603-, 69318-, 12303-, 48000-, 52920-, 5433-, 38554-, 57301-, 58324-, 55063-, 52991-, 59914-, 59921- or 33751-pcDNA/Amp plasmid DNA using the calcium phosphate or calcium chloride co-precipitation methods, DEAE-dextran-mediated transfection, lipofection, or electroporation. Other suitable methods for transfecting host cells can be found 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, N.Y., 1989. The expression of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide is detected by radiolabelling (^{35}S-methionine or ^{35}S-cysteine available from NEN, Boston, Mass., can be used) and immunoprecipitation (Harlow, E. and Lane, D. *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1988) using an HA specific monoclonal antibody. Briefly, the cells are labeled for 8 hours with ^{35}S-methionine (or ^{35}S-cysteine). The culture media are then collected and the cells are lysed using detergents (RIPA buffer, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM Tris, pH 7.5). Both the cell lysate and the culture media are precipitated with an HA specific monoclonal antibody. Precipitated polypeptides are then analyzed by SDS-PAGE.

[0803] Alternatively, DNA containing the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 coding sequence is cloned directly into the polylinker of the pcDNA/Amp vector using the appropriate restriction sites. The resulting plasmid is transfected into COS cells in the manner described above, and the expression of the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide is detected by radiolabelling and immunoprecipitation using a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 specific monoclonal antibody.

Example 4

TaqMan Analysis of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751

[0804] Human 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 expression was measured by TaqMan® quantitative PCR (Perkin Elmer Applied Biosystems) in cDNA prepared from a variety of normal and diseased (e.g., cancerous) human tissues or cell lines.

[0805] Probes were designed by PrimerExpress software (PE Biosystems) based on the sequence of the human 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene. Each human 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene probe was labeled using FAM (6-carboxyfluorescein), and the β2-microglobulin reference probe was labeled with a different fluorescent dye, VIC. The differential labeling of the target gene and internal reference gene thus enabled measurement in same well. Forward and reverse primers and the probes for both β2-microglobulin and target gene were added to the TaqMan® Universal PCR Master Mix (PE Applied Biosystems). Although the final concentration of primer and

probe could vary, each was internally consistent within a given experiment. A typical experiment contained 200 nM of forward and reverse primers plus 100 nM probe for 0-2 microglobulin and 600 nM forward and reverse primers plus 200 nM probe for the target gene. TaqMan matrix experiments were carried out on an ABI PRISM 7700 Sequence Detection System (PE Applied Biosystems). The thermal cycler conditions were as follows: hold for 2 min at 50° C. and 10 min at 95° C., followed by two-step PCR for 40 cycles of 95° C. for 15 sec followed by 60° C. for 1 min.

[0806] The following method was used to quantitatively calculate human 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene expression in the various tissues relative to β-2 microglobulin expression in the same tissue. The threshold cycle (Ct) value is defined as the cycle at which a statistically significant increase in fluorescence is detected. A lower Ct value is indicative of a higher mRNA concentration. The Ct value of the human 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene is normalized by subtracting the Ct value of the β-2 microglobulin gene to obtain a $_{\Delta}$Ct value using the following formula: $_{\Delta}Ct = Ct_{sample} - Ct_{\beta\text{-}2\ microglobulin}$. Expression is then calibrated against a cDNA sample showing a comparatively low level of expression of the human 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 gene. The $_{\Delta}$Ct value for the calibrator sample is then subtracted from $_{\Delta}$Ct for each tissue sample according to the following formula: $_{\Delta\Delta}Ct = {}_{\Delta}Ct\text{-}_{sample} - {}_{\Delta}Ct\text{-}_{calibrator}$. Relative expression is then calculated using the arithmetic formula given by $2^{-\Delta\Delta Ct}$.

Example 5

In Situ Hybridization of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751

[0807] The following describes the tissue distribution of 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 mRNA, as may be determined by in situ hybridization analysis using oligonucleotide probes based on the human G2RF sequence.

[0808] For in situ analysis, various tissues, e.g. tissues obtained from brain, are first frozen on dry ice. Ten-micrometer-thick sections of the tissues are postfixed with 4% formaldehyde in DEPC treated 1× phosphate-buffered saline at room temperature for 10 minutes before being rinsed twice in DEPC 1× phosphate-buffered saline and once in 0.1 M triethanolamine-HCl (pH 8.0). Following incubation in 0.25% acetic anhydride-0.1 M triethanolamine-HCl for 10 minutes, sections are rinsed in DEPC 2×SSC (1×SSC is 0.15M NaCl plus 0.015M sodium citrate). Tissue is then dehydrated through a series of ethanol washes, incubated in 100% chloroform for 5 minutes, and then rinsed in 100% ethanol for 1 minute and 95% ethanol for 1 minute and allowed to air dry.

[0809] Hybridizations are performed with ^{35}S-radiolabeled (5×10^{7} cpm/ml) cRNA probes. Probes are incubated in the presence of a solution containing 600 mM NaCl, 10 mM Tris (pH 7.5), 1 mM EDTA, 0.01% sheared salmon sperm DNA, 0.01% yeast tRNA, 0.05% yeast total RNA type X1, 1×Denhardt's solution, 50% formamide, 10% dextran sulfate, 100 mM dithiothreitol, 0.1% sodium dodecyl sulfate (SDS), and 0.1% sodium thiosulfate for 18 hours at 55° C.

[0810] After hybridization, slides are washed with 2×SSC. Sections are then sequentially incubated at 37° C. in TNE (a solution containing 10 mM Tris-HCl (pH 7.6), 500 mM NaCl, and 1 mM EDTA), for 10 minutes, in TNE with 10 μg of RNase A per ml for 30 minutes, and finally in TNE for 10 minutes. Slides are then rinsed with 2×SSC at room temperature, washed with 2×SSC at 50° C. for 1 hour, washed with 0.2×SSC at 55° C. for 1 hour, and 0.2×SSC at 60° C. for 1 hour. Sections are then dehydrated rapidly through serial ethanol-0.3 M sodium acetate concentrations before being air dried and exposed to Kodak Biomax MR scientific imaging film for 24 hours and subsequently dipped in NB-2 photoemulsion and exposed at 4° C. for 7 days before being developed and counter stained.

[0811] The contents of all references, patents and published patent applications cited throughout this application are incorporated herein by reference.

EQUIVALENTS

[0812] 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.

SEQUENCE LISTING

```
<160> NUMBER OF SEQ ID NOS: 127

<210> SEQ ID NO 1
<211> LENGTH: 3900
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (138)...(3527)

<400> SEQUENCE: 1

cggcccatct ctctgggtct ctgtccctct ctctctgggt ctctgtcccc gtctctctgg      60

gtctcggtcc ccgtctctct gggtctctgt ccccgtctct ctgggtctct gtccccctcc     120

ctgtgtgccc cgctccc atg tgt cca cag ttc ctc cgg ctc tct gac cga        170
```

-continued

```
Met Cys Pro Gln Phe Leu Arg Leu Ser Asp Arg
 1               5                  10

acg gat cca gct gca gtt tat agt ctg gtc aca cgc aca tgg ggc ttc      218
Thr Asp Pro Ala Ala Val Tyr Ser Leu Val Thr Arg Thr Trp Gly Phe
            15                  20                  25

cgt gcc ccg aac ctg gtg gtg tca gtg ctg ggg gga tcg ggg ggc ccc      266
Arg Ala Pro Asn Leu Val Val Ser Val Leu Gly Gly Ser Gly Gly Pro
        30                  35                  40

gtc ctc cag acc tgg ctg cag gac ctg ctg cgt cgt ggg ctg gtg cgg      314
Val Leu Gln Thr Trp Leu Gln Asp Leu Leu Arg Arg Gly Leu Val Arg
    45                  50                  55

gct gcc cag agc aca gga gcc tgg att gtc act ggg ggt ctg cac acg      362
Ala Ala Gln Ser Thr Gly Ala Trp Ile Val Thr Gly Gly Leu His Thr
 60                  65                  70                  75

ggc atc ggc cgg cat gtt ggt gtg gct gta cgg gac cat cag atg gcc      410
Gly Ile Gly Arg His Val Gly Val Ala Val Arg Asp His Gln Met Ala
                 80                  85                  90

agc act ggg ggc acc aag gtg gtg gcc atg ggt gtg gcc ccc tgg ggt      458
Ser Thr Gly Gly Thr Lys Val Val Ala Met Gly Val Ala Pro Trp Gly
            95                 100                 105

gtg gtc cgg aat aga gac acc ctc atc aac ccc aag ggc tcg ttc cct      506
Val Val Arg Asn Arg Asp Thr Leu Ile Asn Pro Lys Gly Ser Phe Pro
        110                 115                 120

gcg agg tac cgg tgg cgc ggt gac ccg gag gac ggg gtc cag ttt ccc      554
Ala Arg Tyr Arg Trp Arg Gly Asp Pro Glu Asp Gly Val Gln Phe Pro
    125                 130                 135

ctg gac tac aac tac tcg gcc ttc ttc ctg gtg gac gac ggc aca cac      602
Leu Asp Tyr Asn Tyr Ser Ala Phe Phe Leu Val Asp Asp Gly Thr His
140                 145                 150                 155

ggc tgc ctg ggg ggc gag aac cgc ttc cgc ttg cgc ctg gag tcc tac      650
Gly Cys Leu Gly Gly Glu Asn Arg Phe Arg Leu Arg Leu Glu Ser Tyr
                160                 165                 170

atc tca cag cag aag acg ggc gtg gga ggg act gga att gac atc cct      698
Ile Ser Gln Gln Lys Thr Gly Val Gly Gly Thr Gly Ile Asp Ile Pro
            175                 180                 185

gtc ctg ctc ctc ctg att gat ggt gat gag aag atg ttg acg cga ata      746
Val Leu Leu Leu Leu Ile Asp Gly Asp Glu Lys Met Leu Thr Arg Ile
        190                 195                 200

gag aac gcc acc cag gct cag ctc cca tgt ctc ctc gtg gct ggc tca      794
Glu Asn Ala Thr Gln Ala Gln Leu Pro Cys Leu Leu Val Ala Gly Ser
    205                 210                 215

ggg gga gct gcg gac tgc ctg gcg gag acc ctg gaa gac act ctg gcc      842
Gly Gly Ala Ala Asp Cys Leu Ala Glu Thr Leu Glu Asp Thr Leu Ala
220                 225                 230                 235

cca ggg agt ggg gga gcc agg caa ggc gaa gcc cga gat cga atc agg      890
Pro Gly Ser Gly Gly Ala Arg Gln Gly Glu Ala Arg Asp Arg Ile Arg
                240                 245                 250

cgt ttc ttt ccc aaa ggg gac ctt gag gtc ctg cag gcc cag gtg gag      938
Arg Phe Phe Pro Lys Gly Asp Leu Glu Val Leu Gln Ala Gln Val Glu
            255                 260                 265

agg att atg acc cgg aag gag ctc ctg aca gtc tat tct tct gag gat      986
Arg Ile Met Thr Arg Lys Glu Leu Leu Thr Val Tyr Ser Ser Glu Asp
        270                 275                 280

ggg tct gag gaa ttc gag acc ata gtt ttg aag gcc ctt gtg aag gcc     1034
Gly Ser Glu Glu Phe Glu Thr Ile Val Leu Lys Ala Leu Val Lys Ala
    285                 290                 295

tgt ggg agc tcg gag gcc tca gcc tac ctg gat gag ctg cgt ttg gct     1082
```

-continued

```
Cys Gly Ser Ser Glu Ala Ser Ala Tyr Leu Asp Glu Leu Arg Leu Ala
300                 305                 310                 315

gtg gct tgg aac cgc gtg gac att gca cag agt gaa ctc ttt cgg ggg     1130
Val Ala Trp Asn Arg Val Asp Ile Ala Gln Ser Glu Leu Phe Arg Gly
                320                 325                 330

gac atc caa tgg cgg tcc ttc cat ctc gaa gct tcc ctc atg gac gcc     1178
Asp Ile Gln Trp Arg Ser Phe His Leu Glu Ala Ser Leu Met Asp Ala
            335                 340                 345

ctg ctg aat gac cgg cct gag ttc gtg cgc ttg ctc att tcc cac ggc     1226
Leu Leu Asn Asp Arg Pro Glu Phe Val Arg Leu Leu Ile Ser His Gly
        350                 355                 360

ctc agc ctg ggc cac ttc ctg acc ccg atg cgc ctg gcc caa ctc tac     1274
Leu Ser Leu Gly His Phe Leu Thr Pro Met Arg Leu Ala Gln Leu Tyr
    365                 370                 375

agc gcg gcg ccc tcc aac tcg ctc atc cgc aac ctt ttg gac cag gcg     1322
Ser Ala Ala Pro Ser Asn Ser Leu Ile Arg Asn Leu Leu Asp Gln Ala
380                 385                 390                 395

tcc cac agc gca ggc acc aaa gcc cca gcc cta aaa ggg gga gct gcg     1370
Ser His Ser Ala Gly Thr Lys Ala Pro Ala Leu Lys Gly Gly Ala Ala
                400                 405                 410

gag ctc cgg ccc cct gac gtg ggg cat gtg ctg agg atg ctg ctg ggg     1418
Glu Leu Arg Pro Pro Asp Val Gly His Val Leu Arg Met Leu Leu Gly
            415                 420                 425

aag atg tgc gcg ccg agg tac ccc tcc ggg ggc gcc tgg gac cct cac     1466
Lys Met Cys Ala Pro Arg Tyr Pro Ser Gly Gly Ala Trp Asp Pro His
        430                 435                 440

cca ggc cag ggc ttc ggg gag agc atg tat ctg ctc tcg gac aag gcc     1514
Pro Gly Gln Gly Phe Gly Glu Ser Met Tyr Leu Leu Ser Asp Lys Ala
    445                 450                 455

acc tcg ccg ctc tcg ctg gat gct ggc ctc ggg cag gcc ccc tgg agc     1562
Thr Ser Pro Leu Ser Leu Asp Ala Gly Leu Gly Gln Ala Pro Trp Ser
460                 465                 470                 475

gac ctg ctt ctt tgg gca ctg ttg ctg aac agg gca cag atg gcc atg     1610
Asp Leu Leu Leu Trp Ala Leu Leu Leu Asn Arg Ala Gln Met Ala Met
                480                 485                 490

tac ttc tgg gag atg ggt tcc aat gca gtt tcc tca gct ctt ggg gcc     1658
Tyr Phe Trp Glu Met Gly Ser Asn Ala Val Ser Ser Ala Leu Gly Ala
            495                 500                 505

tgt ttg ctg ctc cgg gtg atg gca cgc ctg gag cct gac gct gag gag     1706
Cys Leu Leu Leu Arg Val Met Ala Arg Leu Glu Pro Asp Ala Glu Glu
        510                 515                 520

gca gca cgg agg aaa gac ctg gcg ttc aag ttt gag ggg atg ggc gtt     1754
Ala Ala Arg Arg Lys Asp Leu Ala Phe Lys Phe Glu Gly Met Gly Val
    525                 530                 535

gac ctc ttt ggc gag tgc tat cgc agc agt gag gtg agg gct gcc cgc     1802
Asp Leu Phe Gly Glu Cys Tyr Arg Ser Ser Glu Val Arg Ala Ala Arg
540                 545                 550                 555

ctc ctc ctc cgt cgc tgc ccg ctc tgg ggg gat gcc act tgc ctc cag     1850
Leu Leu Leu Arg Arg Cys Pro Leu Trp Gly Asp Ala Thr Cys Leu Gln
                560                 565                 570

ctg gcc atg caa gct gac gcc cgt gcc ttc ttt gcc cag gat ggg gta     1898
Leu Ala Met Gln Ala Asp Ala Arg Ala Phe Phe Ala Gln Asp Gly Val
            575                 580                 585

cag tct ctg ctg aca cag aag tgg tgg gga gat atg gcc agc act aca     1946
Gln Ser Leu Leu Thr Gln Lys Trp Trp Gly Asp Met Ala Ser Thr Thr
        590                 595                 600

ccc atc tgg gcc ctg gtt ctc gcc ttc ttt tgc cct cca ctc atc tac     1994
```

-continued

```
Pro Ile Trp Ala Leu Val Leu Ala Phe Phe Cys Pro Pro Leu Ile Tyr
    605                 610                 615

acc cgc ctc atc acc ttc agg aaa tca gaa gag gag ccc aca cgg gag     2042
Thr Arg Leu Ile Thr Phe Arg Lys Ser Glu Glu Glu Pro Thr Arg Glu
620                 625                 630                 635

gag cta gag ttt gac atg gat agt gtc att aat ggg gaa ggg cct gtc     2090
Glu Leu Glu Phe Asp Met Asp Ser Val Ile Asn Gly Glu Gly Pro Val
                640                 645                 650

ggg acg gcg gac cca gcc gag aag acg ccg ctg ggg gtc ccg cgc cag     2138
Gly Thr Ala Asp Pro Ala Glu Lys Thr Pro Leu Gly Val Pro Arg Gln
            655                 660                 665

tcg ggc cgt ccg ggt tgc tgc ggg ggc cgc tgc ggg ggg cgc cgg tgc     2186
Ser Gly Arg Pro Gly Cys Cys Gly Gly Arg Cys Gly Gly Arg Arg Cys
        670                 675                 680

cta cgc cgc tgg ttc cac ttc tgg ggc gcg ccg gtg acc atc ttc atg     2234
Leu Arg Arg Trp Phe His Phe Trp Gly Ala Pro Val Thr Ile Phe Met
    685                 690                 695

ggc aac gtg gtc agc tac ctg ctg ttc ctg ctg ctt ttc tcg cgg gtg     2282
Gly Asn Val Val Ser Tyr Leu Leu Phe Leu Leu Leu Phe Ser Arg Val
700                 705                 710                 715

ctg ctc gtg gat ttc cag ccg gcg ccg ccc ggc tcc ctg gag ctg ctg     2330
Leu Leu Val Asp Phe Gln Pro Ala Pro Pro Gly Ser Leu Glu Leu Leu
                720                 725                 730

ctc tat ttc tgg gct ttc acg ctg ctg tgc gag gaa ctg cgc cag ggc     2378
Leu Tyr Phe Trp Ala Phe Thr Leu Leu Cys Glu Glu Leu Arg Gln Gly
            735                 740                 745

ctg agc gga ggc ggg ggc agc ctc gcc agc ggg ggc ccc ggg cct ggc     2426
Leu Ser Gly Gly Gly Gly Ser Leu Ala Ser Gly Gly Pro Gly Pro Gly
        750                 755                 760

cat gcc tca ctg agc cag cgc ctg cgc ctc tac ctc gcc gac agc tgg     2474
His Ala Ser Leu Ser Gln Arg Leu Arg Leu Tyr Leu Ala Asp Ser Trp
    765                 770                 775

aac cag tgc gac cta gtg gct ctc acc tgc ttc ctc ctg ggc gtg ggc     2522
Asn Gln Cys Asp Leu Val Ala Leu Thr Cys Phe Leu Leu Gly Val Gly
780                 785                 790                 795

tgc cgg ctg acc ccg ggt ttg tac cac ctg ggc cgc act gtc ctc tgc     2570
Cys Arg Leu Thr Pro Gly Leu Tyr His Leu Gly Arg Thr Val Leu Cys
                800                 805                 810

atc gac ttc atg gtt ttc acg gtg cgg ctg ctt cac atc ttc acg gtc     2618
Ile Asp Phe Met Val Phe Thr Val Arg Leu Leu His Ile Phe Thr Val
            815                 820                 825

aac aaa cag ctg ggg ccc aag atc gtc atc gtg agc aag atg atg aag     2666
Asn Lys Gln Leu Gly Pro Lys Ile Val Ile Val Ser Lys Met Met Lys
        830                 835                 840

gac gtg ttc ttc ttc ctc ttc ttc ctc ggc gtg tgg ctg gta gcc tat     2714
Asp Val Phe Phe Phe Leu Phe Phe Leu Gly Val Trp Leu Val Ala Tyr
    845                 850                 855

ggc gtg gcc acg gag ggg ctc ctg agg cca cgg gac agt gac ttc cca     2762
Gly Val Ala Thr Glu Gly Leu Leu Arg Pro Arg Asp Ser Asp Phe Pro
860                 865                 870                 875

agt atc ctg cgc cgc gtc ttc tac cgt ccc tac ctg cag atc ttc ggg     2810
Ser Ile Leu Arg Arg Val Phe Tyr Arg Pro Tyr Leu Gln Ile Phe Gly
                880                 885                 890

cag att ccc cag gag gac atg gac gtg gcc ctc atg gag cac agc aac     2858
Gln Ile Pro Gln Glu Asp Met Asp Val Ala Leu Met Glu His Ser Asn
            895                 900                 905

tgc tcg tcg gag ccc ggc ttc tgg gca cac cct cct ggg gcc cag gcg     2906
```

-continued

```
Cys Ser Ser Glu Pro Gly Phe Trp Ala His Pro Pro Gly Ala Gln Ala
        910                 915                 920

ggc acc tgc gtc tcc cag tat gcc aac tgg ctg gtg gtg ctg ctc ctc     2954
Gly Thr Cys Val Ser Gln Tyr Ala Asn Trp Leu Val Val Leu Leu Leu
    925                 930                 935

gtc atc ttc ctg ctc gtg gcc aac atc ctg ctg gtc aac ttg ctc att     3002
Val Ile Phe Leu Leu Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile
940                 945                 950                 955

gcc atg ttc agt tac aca ttc ggc aaa gta cag ggc aac agc gat ctc     3050
Ala Met Phe Ser Tyr Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu
                960                 965                 970

tac tgg aag gcg cag cgt tac cgc ctc atc cgg gaa ttc cac tct cgg     3098
Tyr Trp Lys Ala Gln Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg
            975                 980                 985

ccc gcg ctg gcc ccg ccc ttt atc gtc atc tcc cac ttg cgc ctc ctg     3146
Pro Ala Leu Ala Pro Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu
        990                 995                 1000

ctc agg caa ttg tgc agg cga ccc cgg agc ccc cag ccg tcc tcc ccg     3194
Leu Arg Gln Leu Cys Arg Arg Pro Arg Ser Pro Gln Pro Ser Ser Pro
    1005                1010                1015

gcc ctc gag cat ttc cgg gtt tac ctt tct aag gaa gcc gag cgg aag     3242
Ala Leu Glu His Phe Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg Lys
1020                1025                1030                1035

ctg cta acg tgg gaa tcg gtg cat aag gag aac ttt ctg ctg gca cgc     3290
Leu Leu Thr Trp Glu Ser Val His Lys Glu Asn Phe Leu Leu Ala Arg
                1040                1045                1050

gct agg gac aag cgg gag agc gac tcc gag cgt ctg aag cgc acg tcc     3338
Ala Arg Asp Lys Arg Glu Ser Asp Ser Glu Arg Leu Lys Arg Thr Ser
            1055                1060                1065

cag aag gtg gac ttg gca ctg aaa cag ctg gga cac atc cgc gag tac     3386
Gln Lys Val Asp Leu Ala Leu Lys Gln Leu Gly His Ile Arg Glu Tyr
        1070                1075                1080

gaa cag cgc ctg aaa gtg ctg gag cgg gag gtc cag cag tgt agc cgc     3434
Glu Gln Arg Leu Lys Val Leu Glu Arg Glu Val Gln Gln Cys Ser Arg
    1085                1090                1095

gtc ctg ggg tgg gtg gcc gag gcc ctg agc cgc tct gcc ttg ctg ccc     3482
Val Leu Gly Trp Val Ala Glu Ala Leu Ser Arg Ser Ala Leu Leu Pro
1100                1105                1110                1115

cca ggt ggg ccg cca ccc cct gac ctg cct ggg tcc aaa gac tga         3527
Pro Gly Gly Pro Pro Pro Pro Asp Leu Pro Gly Ser Lys Asp  *
                1120                1125

gccctgctgg cggacttcaa ggagaagccc ccacagggga ttttgctcct agagtaaggc   3587

tcatctgggc ctcggccccc gcacctggtg gccttgtcct tgaggtgagc cccatgtcca   3647

tctgggccac tgtcaggacc acctttggga gtgtcatcct tacaaaccac agcatgcccg   3707

gctcctccca gaaccagtcc cagcctggga ggatcaaggc ctggatcccg ggccgttatc   3767

catctggagg ctgcagggtc cttggggtaa cagggaccac agacccctca ccactcacag   3827

attcctcaca ctggggaaat aaagccattt cagaggaaaa aaaaaaaaaa aaaarraaaa   3887

aaaaaaaagg cgg                                                      3900
```

<210> SEQ ID NO 2
<211> LENGTH: 1129
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2

-continued

```
Met Cys Pro Gln Phe Leu Arg Leu Ser Asp Arg Thr Asp Pro Ala Ala
 1               5                  10                  15

Val Tyr Ser Leu Val Thr Arg Thr Trp Gly Phe Arg Ala Pro Asn Leu
            20                  25                  30

Val Val Ser Val Leu Gly Gly Ser Gly Gly Pro Val Leu Gln Thr Trp
        35                  40                  45

Leu Gln Asp Leu Leu Arg Arg Gly Leu Val Arg Ala Ala Gln Ser Thr
    50                  55                  60

Gly Ala Trp Ile Val Thr Gly Gly Leu His Thr Gly Ile Gly Arg His
65                  70                  75                  80

Val Gly Val Ala Val Arg Asp His Gln Met Ala Ser Thr Gly Gly Thr
                85                  90                  95

Lys Val Val Ala Met Gly Val Ala Pro Trp Gly Val Val Arg Asn Arg
            100                 105                 110

Asp Thr Leu Ile Asn Pro Lys Gly Ser Phe Pro Ala Arg Tyr Arg Trp
        115                 120                 125

Arg Gly Asp Pro Glu Asp Gly Val Gln Phe Pro Leu Asp Tyr Asn Tyr
    130                 135                 140

Ser Ala Phe Phe Leu Val Asp Asp Gly Thr His Gly Cys Leu Gly Gly
145                 150                 155                 160

Glu Asn Arg Phe Arg Leu Arg Leu Glu Ser Tyr Ile Ser Gln Gln Lys
                165                 170                 175

Thr Gly Val Gly Gly Thr Gly Ile Asp Ile Pro Val Leu Leu Leu Leu
            180                 185                 190

Ile Asp Gly Asp Glu Lys Met Leu Thr Arg Ile Glu Asn Ala Thr Gln
        195                 200                 205

Ala Gln Leu Pro Cys Leu Leu Val Ala Gly Ser Gly Gly Ala Ala Asp
    210                 215                 220

Cys Leu Ala Glu Thr Leu Glu Asp Thr Leu Ala Pro Gly Ser Gly Gly
225                 230                 235                 240

Ala Arg Gln Gly Glu Ala Arg Asp Arg Ile Arg Arg Phe Phe Pro Lys
                245                 250                 255

Gly Asp Leu Glu Val Leu Gln Ala Gln Val Glu Arg Ile Met Thr Arg
            260                 265                 270

Lys Glu Leu Leu Thr Val Tyr Ser Ser Glu Asp Gly Ser Glu Glu Phe
        275                 280                 285

Glu Thr Ile Val Leu Lys Ala Leu Val Lys Ala Cys Gly Ser Ser Glu
    290                 295                 300

Ala Ser Ala Tyr Leu Asp Glu Leu Arg Leu Ala Val Ala Trp Asn Arg
305                 310                 315                 320

Val Asp Ile Ala Gln Ser Glu Leu Phe Arg Gly Asp Ile Gln Trp Arg
                325                 330                 335

Ser Phe His Leu Glu Ala Ser Leu Met Asp Ala Leu Leu Asn Asp Arg
            340                 345                 350

Pro Glu Phe Val Arg Leu Leu Ile Ser His Gly Leu Ser Leu Gly His
        355                 360                 365

Phe Leu Thr Pro Met Arg Leu Ala Gln Leu Tyr Ser Ala Ala Pro Ser
    370                 375                 380

Asn Ser Leu Ile Arg Asn Leu Leu Asp Gln Ala Ser His Ser Ala Gly
385                 390                 395                 400
```

-continued

```
Thr Lys Ala Pro Ala Leu Lys Gly Gly Ala Ala Glu Leu Arg Pro Pro
                405                 410                 415

Asp Val Gly His Val Leu Arg Met Leu Leu Gly Lys Met Cys Ala Pro
            420                 425                 430

Arg Tyr Pro Ser Gly Gly Ala Trp Asp Pro His Pro Gly Gln Gly Phe
        435                 440                 445

Gly Glu Ser Met Tyr Leu Leu Ser Asp Lys Ala Thr Ser Pro Leu Ser
    450                 455                 460

Leu Asp Ala Gly Leu Gly Gln Ala Pro Trp Ser Asp Leu Leu Leu Trp
465                 470                 475                 480

Ala Leu Leu Leu Asn Arg Ala Gln Met Ala Met Tyr Phe Trp Glu Met
                485                 490                 495

Gly Ser Asn Ala Val Ser Ser Ala Leu Gly Ala Cys Leu Leu Leu Arg
            500                 505                 510

Val Met Ala Arg Leu Glu Pro Asp Ala Glu Glu Ala Ala Arg Arg Lys
        515                 520                 525

Asp Leu Ala Phe Lys Phe Glu Gly Met Gly Val Asp Leu Phe Gly Glu
    530                 535                 540

Cys Tyr Arg Ser Ser Glu Val Arg Ala Ala Arg Leu Leu Leu Arg Arg
545                 550                 555                 560

Cys Pro Leu Trp Gly Asp Ala Thr Cys Leu Gln Leu Ala Met Gln Ala
                565                 570                 575

Asp Ala Arg Ala Phe Phe Ala Gln Asp Gly Val Gln Ser Leu Leu Thr
            580                 585                 590

Gln Lys Trp Trp Gly Asp Met Ala Ser Thr Thr Pro Ile Trp Ala Leu
        595                 600                 605

Val Leu Ala Phe Phe Cys Pro Pro Leu Ile Tyr Thr Arg Leu Ile Thr
    610                 615                 620

Phe Arg Lys Ser Glu Glu Glu Pro Thr Arg Glu Glu Leu Glu Phe Asp
625                 630                 635                 640

Met Asp Ser Val Ile Asn Gly Glu Gly Pro Val Gly Thr Ala Asp Pro
                645                 650                 655

Ala Glu Lys Thr Pro Leu Gly Val Pro Arg Gln Ser Gly Arg Pro Gly
            660                 665                 670

Cys Cys Gly Gly Arg Cys Gly Gly Arg Arg Cys Leu Arg Arg Trp Phe
        675                 680                 685

His Phe Trp Gly Ala Pro Val Thr Ile Phe Met Gly Asn Val Val Ser
    690                 695                 700

Tyr Leu Leu Phe Leu Leu Leu Phe Ser Arg Val Leu Leu Val Asp Phe
705                 710                 715                 720

Gln Pro Ala Pro Pro Gly Ser Leu Glu Leu Leu Leu Tyr Phe Trp Ala
                725                 730                 735

Phe Thr Leu Leu Cys Glu Glu Leu Arg Gln Gly Leu Ser Gly Gly Gly
            740                 745                 750

Gly Ser Leu Ala Ser Gly Gly Pro Gly Pro Gly His Ala Ser Leu Ser
        755                 760                 765

Gln Arg Leu Arg Leu Tyr Leu Ala Asp Ser Trp Asn Gln Cys Asp Leu
    770                 775                 780

Val Ala Leu Thr Cys Phe Leu Leu Gly Val Gly Cys Arg Leu Thr Pro
785                 790                 795                 800

Gly Leu Tyr His Leu Gly Arg Thr Val Leu Cys Ile Asp Phe Met Val
```

-continued

```
                805                 810                 815

Phe Thr Val Arg Leu Leu His Ile Phe Thr Val Asn Lys Gln Leu Gly
            820                 825                 830

Pro Lys Ile Val Ile Val Ser Lys Met Met Lys Asp Val Phe Phe Phe
        835                 840                 845

Leu Phe Phe Leu Gly Val Trp Leu Val Ala Tyr Gly Val Ala Thr Glu
    850                 855                 860

Gly Leu Leu Arg Pro Arg Asp Ser Asp Phe Pro Ser Ile Leu Arg Arg
865                 870                 875                 880

Val Phe Tyr Arg Pro Tyr Leu Gln Ile Phe Gly Gln Ile Pro Gln Glu
                885                 890                 895

Asp Met Asp Val Ala Leu Met Glu His Ser Asn Cys Ser Ser Glu Pro
            900                 905                 910

Gly Phe Trp Ala His Pro Pro Gly Ala Gln Ala Gly Thr Cys Val Ser
        915                 920                 925

Gln Tyr Ala Asn Trp Leu Val Val Leu Leu Leu Val Ile Phe Leu Leu
    930                 935                 940

Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Met Phe Ser Tyr
945                 950                 955                 960

Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu Tyr Trp Lys Ala Gln
                965                 970                 975

Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg Pro Ala Leu Ala Pro
            980                 985                 990

Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu Leu Arg Gln Leu Cys
        995                 1000                1005

Arg Arg Pro Arg Ser Pro Gln Pro Ser Ser Pro Ala Leu Glu His Phe
    1010                1015                1020

Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg Lys Leu Leu Thr Trp Glu
1025                1030                1035                1040

Ser Val His Lys Glu Asn Phe Leu Leu Ala Arg Ala Arg Asp Lys Arg
                1045                1050                1055

Glu Ser Asp Ser Glu Arg Leu Lys Arg Thr Ser Gln Lys Val Asp Leu
            1060                1065                1070

Ala Leu Lys Gln Leu Gly His Ile Arg Glu Tyr Glu Gln Arg Leu Lys
        1075                1080                1085

Val Leu Glu Arg Glu Val Gln Gln Cys Ser Arg Val Leu Gly Trp Val
    1090                1095                1100

Ala Glu Ala Leu Ser Arg Ser Ala Leu Leu Pro Pro Gly Gly Pro Pro
1105                1110                1115                1120

Pro Pro Asp Leu Pro Gly Ser Lys Asp
                1125


<210> SEQ ID NO 3
<211> LENGTH: 3390
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(3390)

<400> SEQUENCE: 3

atg tgt cca cag ttc ctc cgg ctc tct gac cga acg gat cca gct gca       48
Met Cys Pro Gln Phe Leu Arg Leu Ser Asp Arg Thr Asp Pro Ala Ala
 1               5                  10                  15
```

-continued

```
gtt tat agt ctg gtc aca cgc aca tgg ggc ttc cgt gcc ccg aac ctg       96
Val Tyr Ser Leu Val Thr Arg Thr Trp Gly Phe Arg Ala Pro Asn Leu
            20                  25                  30

gtg gtg tca gtg ctg ggg gga tcg ggg ggc ccc gtc ctc cag acc tgg      144
Val Val Ser Val Leu Gly Gly Ser Gly Gly Pro Val Leu Gln Thr Trp
        35                  40                  45

ctg cag gac ctg ctg cgt cgt ggg ctg gtg cgg gct gcc cag agc aca      192
Leu Gln Asp Leu Leu Arg Arg Gly Leu Val Arg Ala Ala Gln Ser Thr
    50                  55                  60

gga gcc tgg att gtc act ggg ggt ctg cac acg ggc atc ggc cgg cat      240
Gly Ala Trp Ile Val Thr Gly Gly Leu His Thr Gly Ile Gly Arg His
65                  70                  75                  80

gtt ggt gtg gct gta cgg gac cat cag atg gcc agc act ggg ggc acc      288
Val Gly Val Ala Val Arg Asp His Gln Met Ala Ser Thr Gly Gly Thr
                85                  90                  95

aag gtg gtg gcc atg ggt gtg gcc ccc tgg ggt gtg gtc cgg aat aga      336
Lys Val Val Ala Met Gly Val Ala Pro Trp Gly Val Val Arg Asn Arg
            100                 105                 110

gac acc ctc atc aac ccc aag ggc tcg ttc cct gcg agg tac cgg tgg      384
Asp Thr Leu Ile Asn Pro Lys Gly Ser Phe Pro Ala Arg Tyr Arg Trp
        115                 120                 125

cgc ggt gac ccg gag gac ggg gtc cag ttt ccc ctg gac tac aac tac      432
Arg Gly Asp Pro Glu Asp Gly Val Gln Phe Pro Leu Asp Tyr Asn Tyr
    130                 135                 140

tcg gcc ttc ttc ctg gtg gac gac ggc aca cac ggc tgc ctg ggg ggc      480
Ser Ala Phe Phe Leu Val Asp Asp Gly Thr His Gly Cys Leu Gly Gly
145                 150                 155                 160

gag aac cgc ttc cgc ttg cgc ctg gag tcc tac atc tca cag cag aag      528
Glu Asn Arg Phe Arg Leu Arg Leu Glu Ser Tyr Ile Ser Gln Gln Lys
                165                 170                 175

acg ggc gtg gga ggg act gga att gac atc cct gtc ctg ctc ctc ctg      576
Thr Gly Val Gly Gly Thr Gly Ile Asp Ile Pro Val Leu Leu Leu Leu
            180                 185                 190

att gat ggt gat gag aag atg ttg acg cga ata gag aac gcc acc cag      624
Ile Asp Gly Asp Glu Lys Met Leu Thr Arg Ile Glu Asn Ala Thr Gln
        195                 200                 205

gct cag ctc cca tgt ctc ctc gtg gct ggc tca ggg gga gct gcg gac      672
Ala Gln Leu Pro Cys Leu Leu Val Ala Gly Ser Gly Gly Ala Ala Asp
    210                 215                 220

tgc ctg gcg gag acc ctg gaa gac act ctg gcc cca ggg agt ggg gga      720
Cys Leu Ala Glu Thr Leu Glu Asp Thr Leu Ala Pro Gly Ser Gly Gly
225                 230                 235                 240

gcc agg caa ggc gaa gcc cga gat cga atc agg cgt ttc ttt ccc aaa      768
Ala Arg Gln Gly Glu Ala Arg Asp Arg Ile Arg Arg Phe Phe Pro Lys
                245                 250                 255

ggg gac ctt gag gtc ctg cag gcc cag gtg gag agg att atg acc cgg      816
Gly Asp Leu Glu Val Leu Gln Ala Gln Val Glu Arg Ile Met Thr Arg
            260                 265                 270

aag gag ctc ctg aca gtc tat tct tct gag gat ggg tct gag gaa ttc      864
Lys Glu Leu Leu Thr Val Tyr Ser Ser Glu Asp Gly Ser Glu Glu Phe
        275                 280                 285

gag acc ata gtt ttg aag gcc ctt gtg aag gcc tgt ggg agc tcg gag      912
Glu Thr Ile Val Leu Lys Ala Leu Val Lys Ala Cys Gly Ser Ser Glu
    290                 295                 300

gcc tca gcc tac ctg gat gag ctg cgt ttg gct gtg gct tgg aac cgc      960
Ala Ser Ala Tyr Leu Asp Glu Leu Arg Leu Ala Val Ala Trp Asn Arg
305                 310                 315                 320
```

-continued

```
gtg gac att gca cag agt gaa ctc ttt cgg ggg gac atc caa tgg cgg     1008
Val Asp Ile Ala Gln Ser Glu Leu Phe Arg Gly Asp Ile Gln Trp Arg
                325                 330                 335

tcc ttc cat ctc gaa gct tcc ctc atg gac gcc ctg ctg aat gac cgg     1056
Ser Phe His Leu Glu Ala Ser Leu Met Asp Ala Leu Leu Asn Asp Arg
            340                 345                 350

cct gag ttc gtg cgc ttg ctc att tcc cac ggc ctc agc ctg ggc cac     1104
Pro Glu Phe Val Arg Leu Leu Ile Ser His Gly Leu Ser Leu Gly His
        355                 360                 365

ttc ctg acc ccg atg cgc ctg gcc caa ctc tac agc gcg gcg ccc tcc     1152
Phe Leu Thr Pro Met Arg Leu Ala Gln Leu Tyr Ser Ala Ala Pro Ser
    370                 375                 380

aac tcg ctc atc cgc aac ctt ttg gac cag gcg tcc cac agc gca ggc     1200
Asn Ser Leu Ile Arg Asn Leu Leu Asp Gln Ala Ser His Ser Ala Gly
385                 390                 395                 400

acc aaa gcc cca gcc cta aaa ggg gga gct gcg gag ctc cgg ccc cct     1248
Thr Lys Ala Pro Ala Leu Lys Gly Gly Ala Ala Glu Leu Arg Pro Pro
                405                 410                 415

gac gtg ggg cat gtg ctg agg atg ctg ctg ggg aag atg tgc gcg ccg     1296
Asp Val Gly His Val Leu Arg Met Leu Leu Gly Lys Met Cys Ala Pro
            420                 425                 430

agg tac ccc tcc ggg ggc gcc tgg gac cct cac cca ggc cag ggc ttc     1344
Arg Tyr Pro Ser Gly Gly Ala Trp Asp Pro His Pro Gly Gln Gly Phe
        435                 440                 445

ggg gag agc atg tat ctg ctc tcg gac aag gcc acc tcg ccg ctc tcg     1392
Gly Glu Ser Met Tyr Leu Leu Ser Asp Lys Ala Thr Ser Pro Leu Ser
    450                 455                 460

ctg gat gct ggc ctc ggg cag gcc ccc tgg agc gac ctg ctt ctt tgg     1440
Leu Asp Ala Gly Leu Gly Gln Ala Pro Trp Ser Asp Leu Leu Leu Trp
465                 470                 475                 480

gca ctg ttg ctg aac agg gca cag atg gcc atg tac ttc tgg gag atg     1488
Ala Leu Leu Leu Asn Arg Ala Gln Met Ala Met Tyr Phe Trp Glu Met
                485                 490                 495

ggt tcc aat gca gtt tcc tca gct ctt ggg gcc tgt ttg ctg ctc cgg     1536
Gly Ser Asn Ala Val Ser Ser Ala Leu Gly Ala Cys Leu Leu Leu Arg
            500                 505                 510

gtg atg gca cgc ctg gag cct gac gct gag gag gca gca cgg agg aaa     1584
Val Met Ala Arg Leu Glu Pro Asp Ala Glu Glu Ala Ala Arg Arg Lys
        515                 520                 525

gac ctg gcg ttc aag ttt gag ggg atg ggc gtt gac ctc ttt ggc gag     1632
Asp Leu Ala Phe Lys Phe Glu Gly Met Gly Val Asp Leu Phe Gly Glu
    530                 535                 540

tgc tat cgc agc agt gag gtg agg gct gcc cgc ctc ctc ctc cgt cgc     1680
Cys Tyr Arg Ser Ser Glu Val Arg Ala Ala Arg Leu Leu Leu Arg Arg
545                 550                 555                 560

tgc ccg ctc tgg ggg gat gcc act tgc ctc cag ctg gcc atg caa gct     1728
Cys Pro Leu Trp Gly Asp Ala Thr Cys Leu Gln Leu Ala Met Gln Ala
                565                 570                 575

gac gcc cgt gcc ttc ttt gcc cag gat ggg gta cag tct ctg ctg aca     1776
Asp Ala Arg Ala Phe Phe Ala Gln Asp Gly Val Gln Ser Leu Leu Thr
            580                 585                 590

cag aag tgg tgg gga gat atg gcc agc act aca ccc atc tgg gcc ctg     1824
Gln Lys Trp Trp Gly Asp Met Ala Ser Thr Thr Pro Ile Trp Ala Leu
        595                 600                 605

gtt ctc gcc ttc ttt tgc cct cca ctc atc tac acc cgc ctc atc acc     1872
Val Leu Ala Phe Phe Cys Pro Pro Leu Ile Tyr Thr Arg Leu Ile Thr
    610                 615                 620
```

-continued

```
ttc agg aaa tca gaa gag gag ccc aca cgg gag gag cta gag ttt gac     1920
Phe Arg Lys Ser Glu Glu Glu Pro Thr Arg Glu Glu Leu Glu Phe Asp
625                 630                 635                 640

atg gat agt gtc att aat ggg gaa ggg cct gtc ggg acg gcg gac cca     1968
Met Asp Ser Val Ile Asn Gly Glu Gly Pro Val Gly Thr Ala Asp Pro
                645                 650                 655

gcc gag aag acg ccg ctg ggg gtc ccg cgc cag tcg ggc cgt ccg ggt     2016
Ala Glu Lys Thr Pro Leu Gly Val Pro Arg Gln Ser Gly Arg Pro Gly
            660                 665                 670

tgc tgc ggg ggc cgc tgc ggg ggg cgc cgg tgc cta cgc cgc tgg ttc     2064
Cys Cys Gly Gly Arg Cys Gly Gly Arg Arg Cys Leu Arg Arg Trp Phe
        675                 680                 685

cac ttc tgg ggc gcg ccg gtg acc atc ttc atg ggc aac gtg gtc agc     2112
His Phe Trp Gly Ala Pro Val Thr Ile Phe Met Gly Asn Val Val Ser
    690                 695                 700

tac ctg ctg ttc ctg ctg ctt ttc tcg cgg gtg ctg ctc gtg gat ttc     2160
Tyr Leu Leu Phe Leu Leu Leu Phe Ser Arg Val Leu Leu Val Asp Phe
705                 710                 715                 720

cag ccg gcg ccg ccc ggc tcc ctg gag ctg ctg ctc tat ttc tgg gct     2208
Gln Pro Ala Pro Pro Gly Ser Leu Glu Leu Leu Leu Tyr Phe Trp Ala
                725                 730                 735

ttc acg ctg ctg tgc gag gaa ctg cgc cag ggc ctg agc gga ggc ggg     2256
Phe Thr Leu Leu Cys Glu Glu Leu Arg Gln Gly Leu Ser Gly Gly Gly
            740                 745                 750

ggc agc ctc gcc agc ggg ggc ccc ggg cct ggc cat gcc tca ctg agc     2304
Gly Ser Leu Ala Ser Gly Gly Pro Gly Pro Gly His Ala Ser Leu Ser
        755                 760                 765

cag cgc ctg cgc ctc tac ctc gcc gac agc tgg aac cag tgc gac cta     2352
Gln Arg Leu Arg Leu Tyr Leu Ala Asp Ser Trp Asn Gln Cys Asp Leu
    770                 775                 780

gtg gct ctc acc tgc ttc ctc ctg ggc gtg ggc tgc cgg ctg acc ccg     2400
Val Ala Leu Thr Cys Phe Leu Leu Gly Val Gly Cys Arg Leu Thr Pro
785                 790                 795                 800

ggt ttg tac cac ctg ggc cgc act gtc ctc tgc atc gac ttc atg gtt     2448
Gly Leu Tyr His Leu Gly Arg Thr Val Leu Cys Ile Asp Phe Met Val
                805                 810                 815

ttc acg gtg cgg ctg ctt cac atc ttc acg gtc aac aaa cag ctg ggg     2496
Phe Thr Val Arg Leu Leu His Ile Phe Thr Val Asn Lys Gln Leu Gly
            820                 825                 830

ccc aag atc gtc atc gtg agc aag atg atg aag gac gtg ttc ttc ttc     2544
Pro Lys Ile Val Ile Val Ser Lys Met Met Lys Asp Val Phe Phe Phe
        835                 840                 845

ctc ttc ttc ctc ggc gtg tgg ctg gta gcc tat ggc gtg gcc acg gag     2592
Leu Phe Phe Leu Gly Val Trp Leu Val Ala Tyr Gly Val Ala Thr Glu
    850                 855                 860

ggg ctc ctg agg cca cgg gac agt gac ttc cca agt atc ctg cgc cgc     2640
Gly Leu Leu Arg Pro Arg Asp Ser Asp Phe Pro Ser Ile Leu Arg Arg
865                 870                 875                 880

gtc ttc tac cgt ccc tac ctg cag atc ttc ggg cag att ccc cag gag     2688
Val Phe Tyr Arg Pro Tyr Leu Gln Ile Phe Gly Gln Ile Pro Gln Glu
                885                 890                 895

gac atg gac gtg gcc ctc atg gag cac agc aac tgc tcg tcg gag ccc     2736
Asp Met Asp Val Ala Leu Met Glu His Ser Asn Cys Ser Ser Glu Pro
            900                 905                 910

ggc ttc tgg gca cac cct cct ggg gcc cag gcg ggc acc tgc gtc tcc     2784
Gly Phe Trp Ala His Pro Pro Gly Ala Gln Ala Gly Thr Cys Val Ser
        915                 920                 925
```

-continued

```
cag tat gcc aac tgg ctg gtg gtg ctg ctc ctc gtc atc ttc ctg ctc     2832
Gln Tyr Ala Asn Trp Leu Val Val Leu Leu Leu Val Ile Phe Leu Leu
    930                 935                 940

gtg gcc aac atc ctg ctg gtc aac ttg ctc att gcc atg ttc agt tac     2880
Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Met Phe Ser Tyr
945                 950                 955                 960

aca ttc ggc aaa gta cag ggc aac agc gat ctc tac tgg aag gcg cag     2928
Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu Tyr Trp Lys Ala Gln
                965                 970                 975

cgt tac cgc ctc atc cgg gaa ttc cac tct cgg ccc gcg ctg gcc ccg     2976
Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg Pro Ala Leu Ala Pro
            980                 985                 990

ccc ttt atc gtc atc tcc cac ttg cgc ctc ctg ctc agg caa ttg tgc     3024
Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu Leu Arg Gln Leu Cys
        995                 1000                1005

agg cga ccc cgg agc ccc cag ccg tcc tcc ccg gcc ctc gag cat ttc     3072
Arg Arg Pro Arg Ser Pro Gln Pro Ser Ser Pro Ala Leu Glu His Phe
    1010                1015                1020

cgg gtt tac ctt tct aag gaa gcc gag cgg aag ctg cta acg tgg gaa     3120
Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg Lys Leu Leu Thr Trp Glu
1025                1030                1035                1040

tcg gtg cat aag gag aac ttt ctg ctg gca cgc gct agg gac aag cgg     3168
Ser Val His Lys Glu Asn Phe Leu Leu Ala Arg Ala Arg Asp Lys Arg
                1045                1050                1055

gag agc gac tcc gag cgt ctg aag cgc acg tcc cag aag gtg gac ttg     3216
Glu Ser Asp Ser Glu Arg Leu Lys Arg Thr Ser Gln Lys Val Asp Leu
            1060                1065                1070

gca ctg aaa cag ctg gga cac atc cgc gag tac gaa cag cgc ctg aaa     3264
Ala Leu Lys Gln Leu Gly His Ile Arg Glu Tyr Glu Gln Arg Leu Lys
        1075                1080                1085

gtg ctg gag cgg gag gtc cag cag tgt agc cgc gtc ctg ggg tgg gtg     3312
Val Leu Glu Arg Glu Val Gln Gln Cys Ser Arg Val Leu Gly Trp Val
    1090                1095                1100

gcc gag gcc ctg agc cgc tct gcc ttg ctg ccc cca ggt ggg ccg cca     3360
Ala Glu Ala Leu Ser Arg Ser Ala Leu Leu Pro Pro Gly Gly Pro Pro
1105                1110                1115                1120

ccc cct gac ctg cct ggg tcc aaa gac tga                             3390
Pro Pro Asp Leu Pro Gly Ser Lys Asp  *
                1125


<210> SEQ ID NO 4
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (2)...(2)
<223> OTHER INFORMATION: Xaa = Any amino acid except P
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (3)...(3)
<223> OTHER INFORMATION: Xaa at position 3 can be S or T
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (4)...(4)
<223> OTHER INFORMATION: Xaa = Any amino acid except P

<400> SEQUENCE: 4

Asn Xaa Xaa Xaa
 1
```

-continued

```
<210> SEQ ID NO 5
<211> LENGTH: 2796
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(2796)

<400> SEQUENCE: 5

atg agc caa agc ccg gcg ttc ggg ccc cgg agg ggc att tct ccc cgg       48
Met Ser Gln Ser Pro Ala Phe Gly Pro Arg Arg Gly Ile Ser Pro Arg
 1               5                  10                  15

ggc gct gcc gga gcc gct gcg cgg cgc aac gag agc cag gac tat ctg       96
Gly Ala Ala Gly Ala Ala Ala Arg Arg Asn Glu Ser Gln Asp Tyr Leu
            20                  25                  30

ctc atg gac tcg gag ctg gga gaa gac ggc tgc ccg caa gcc ccg ctg      144
Leu Met Asp Ser Glu Leu Gly Glu Asp Gly Cys Pro Gln Ala Pro Leu
        35                  40                  45

cct tgc tac ggc tac tac ccc tgc ttc cgg gga tct gac aac aga ctg      192
Pro Cys Tyr Gly Tyr Tyr Pro Cys Phe Arg Gly Ser Asp Asn Arg Leu
    50                  55                  60

gct cac cgg cgg cag aca gtt ctc cgt gag aag ggg aga agg tta gct      240
Ala His Arg Arg Gln Thr Val Leu Arg Glu Lys Gly Arg Arg Leu Ala
65                  70                  75                  80

aat cga gga cca gca tac atg ttt agt gat cgc tcc aca agc cta tct      288
Asn Arg Gly Pro Ala Tyr Met Phe Ser Asp Arg Ser Thr Ser Leu Ser
                85                  90                  95

ata gag gag gaa cgc ttt ttg gat gca gct gaa tat ggt aac atc cca      336
Ile Glu Glu Glu Arg Phe Leu Asp Ala Ala Glu Tyr Gly Asn Ile Pro
            100                 105                 110

gtg gtg cgg aag atg tta gaa gaa tgc cac tca ctc aac gtt aac tgt      384
Val Val Arg Lys Met Leu Glu Glu Cys His Ser Leu Asn Val Asn Cys
        115                 120                 125

gtg gat tac atg ggc cag aat gcc cta cag ttg gca gtg gcc aat gag      432
Val Asp Tyr Met Gly Gln Asn Ala Leu Gln Leu Ala Val Ala Asn Glu
    130                 135                 140

cat ctg gaa att aca gaa ctt ctt ctc aag aaa gaa aac ctc tct cga      480
His Leu Glu Ile Thr Glu Leu Leu Leu Lys Lys Glu Asn Leu Ser Arg
145                 150                 155                 160

att ggg gat gct ttg ctt cta gct att agt aaa ggt tat gtt cgg att      528
Ile Gly Asp Ala Leu Leu Leu Ala Ile Ser Lys Gly Tyr Val Arg Ile
                165                 170                 175

gtg gaa gca att ctc agt cat ccg gct ttt gct gaa ggc aag agg tta      576
Val Glu Ala Ile Leu Ser His Pro Ala Phe Ala Glu Gly Lys Arg Leu
            180                 185                 190

gca acc agc cct agc cag tct gaa ctc cag caa gat gat ttt tat gcc      624
Ala Thr Ser Pro Ser Gln Ser Glu Leu Gln Gln Asp Asp Phe Tyr Ala
        195                 200                 205

tat gat gaa gat ggg aca cgg tcc tcc cat gat gtg act cca atc att      672
Tyr Asp Glu Asp Gly Thr Arg Ser Ser His Asp Val Thr Pro Ile Ile
    210                 215                 220

ctg gct gcc cac tgc cgg gaa tat gaa att gtg cat acc ctc ctg cgg      720
Leu Ala Ala His Cys Arg Glu Tyr Glu Ile Val His Thr Leu Leu Arg
225                 230                 235                 240

aag ggt gct agg att gaa cgg cct cat gat tat ttc tgc aag tgc aat      768
Lys Gly Ala Arg Ile Glu Arg Pro His Asp Tyr Phe Cys Lys Cys Asn
                245                 250                 255

gac tgc aac cag aaa cag aag cat gac tcg ttt agc cac tcc aga tct      816
```

-continued

```
Asp Cys Asn Gln Lys Gln Lys His Asp Ser Phe Ser His Ser Arg Ser
            260                 265                 270

agg att aat gcc tat aaa ggc ctg gca agt ccg gct tac ctg tca ttg      864
Arg Ile Asn Ala Tyr Lys Gly Leu Ala Ser Pro Ala Tyr Leu Ser Leu
        275                 280                 285

tct agt gaa gat cca gtc atg acg gct tta gaa ctt agc aat gaa ctg      912
Ser Ser Glu Asp Pro Val Met Thr Ala Leu Glu Leu Ser Asn Glu Leu
    290                 295                 300

gca gtt ctg gcc aat att gag aaa gag ttc aag aat gac tac aaa aaa      960
Ala Val Leu Ala Asn Ile Glu Lys Glu Phe Lys Asn Asp Tyr Lys Lys
305                 310                 315                 320

ctg tca atg cag tgc aaa gac ttt gtt gtt gga ctc ctt gat ctg tgc     1008
Leu Ser Met Gln Cys Lys Asp Phe Val Val Gly Leu Leu Asp Leu Cys
                325                 330                 335

aga aac act gaa gaa gtc gag gcc att ctg aat ggg gat gtt gaa acg     1056
Arg Asn Thr Glu Glu Val Glu Ala Ile Leu Asn Gly Asp Val Glu Thr
            340                 345                 350

ctc cag agt ggt gat cac ggt cgc cca aat ctc agc cgt tta aaa ctt     1104
Leu Gln Ser Gly Asp His Gly Arg Pro Asn Leu Ser Arg Leu Lys Leu
        355                 360                 365

gcc att aaa tat gaa gta aaa aaa ttt gta gct cat cca aac tgc caa     1152
Ala Ile Lys Tyr Glu Val Lys Lys Phe Val Ala His Pro Asn Cys Gln
    370                 375                 380

cag caa ctt ctc tcc att tgg tat gag aat ctt tct ggt tta cga cag     1200
Gln Gln Leu Leu Ser Ile Trp Tyr Glu Asn Leu Ser Gly Leu Arg Gln
385                 390                 395                 400

cag aca atg gcg gtc aag ttc ctt gtg gcc ctt gct gtt gcc att gga     1248
Gln Thr Met Ala Val Lys Phe Leu Val Ala Leu Ala Val Ala Ile Gly
                405                 410                 415

ctg ccc ttc ctg gct ctc att tac tgg ttt gct cca tgc agc aag atg     1296
Leu Pro Phe Leu Ala Leu Ile Tyr Trp Phe Ala Pro Cys Ser Lys Met
            420                 425                 430

ggg aag ata atg cgt gga cca ttc atg aag ttt gta gca cac gca gcc     1344
Gly Lys Ile Met Arg Gly Pro Phe Met Lys Phe Val Ala His Ala Ala
        435                 440                 445

tcc ttc acc att ttt ctg gga ctg cta gtc atg aat gca gct gac aga     1392
Ser Phe Thr Ile Phe Leu Gly Leu Leu Val Met Asn Ala Ala Asp Arg
    450                 455                 460

ttt gaa ggc aca aaa ctc ctt cct aat gaa acc agc aca gat aat gca     1440
Phe Glu Gly Thr Lys Leu Leu Pro Asn Glu Thr Ser Thr Asp Asn Ala
465                 470                 475                 480

aaa cag ctg ttc agg atg aaa aca tcc tgc ttc tca tgg atg gag atg     1488
Lys Gln Leu Phe Arg Met Lys Thr Ser Cys Phe Ser Trp Met Glu Met
                485                 490                 495

ctc att ata tcc tgg gta ata ggc atg ata tgg gct gaa tgt aaa gaa     1536
Leu Ile Ile Ser Trp Val Ile Gly Met Ile Trp Ala Glu Cys Lys Glu
            500                 505                 510

atc tgg act cag ggc ccc aag gaa tat ttg ttt gag ttg tgg aac atg     1584
Ile Trp Thr Gln Gly Pro Lys Glu Tyr Leu Phe Glu Leu Trp Asn Met
        515                 520                 525

ctt gat ttt ggt atg tta gca att ttc cca gca tca ttc att gcg aga     1632
Leu Asp Phe Gly Met Leu Ala Ile Phe Pro Ala Ser Phe Ile Ala Arg
    530                 535                 540

ttc atg gca ttt tgg cat gct tcc aaa gcc cag agc atc att gac gca     1680
Phe Met Ala Phe Trp His Ala Ser Lys Ala Gln Ser Ile Ile Asp Ala
545                 550                 555                 560

aac gat act ttg aag gac ttg acg aaa gta aca ttg gga gac aat gtg     1728
```

-continued

```
Asn Asp Thr Leu Lys Asp Leu Thr Lys Val Thr Leu Gly Asp Asn Val
                565                 570                 575

aaa tac tac aat ttg gcc agg ata aag tgg gac ccc tct gat cct caa     1776
Lys Tyr Tyr Asn Leu Ala Arg Ile Lys Trp Asp Pro Ser Asp Pro Gln
            580                 585                 590

ata ata tct gaa ggt ctt tat gca att gct gta gtt tta agt ttc tct     1824
Ile Ile Ser Glu Gly Leu Tyr Ala Ile Ala Val Val Leu Ser Phe Ser
        595                 600                 605

agg ata gct tat att tta cca gca aat gaa agc ttt gga cct ctg cag     1872
Arg Ile Ala Tyr Ile Leu Pro Ala Asn Glu Ser Phe Gly Pro Leu Gln
    610                 615                 620

ata tca ctt gga aga aca gtc aaa gac atc ttc aag ttc atg gtc ata     1920
Ile Ser Leu Gly Arg Thr Val Lys Asp Ile Phe Lys Phe Met Val Ile
625                 630                 635                 640

ttc att atg gtg ttt gtg gcc ttt atg att gga atg ttc aat ctc tac     1968
Phe Ile Met Val Phe Val Ala Phe Met Ile Gly Met Phe Asn Leu Tyr
                645                 650                 655

tcc tac tac att ggt gca aaa caa aat gaa gcc ttc aca aca gtt gaa     2016
Ser Tyr Tyr Ile Gly Ala Lys Gln Asn Glu Ala Phe Thr Thr Val Glu
            660                 665                 670

gag agt ttt aag aca ctg ttc tgg gct ata ttt gga ctt tct gaa gtg     2064
Glu Ser Phe Lys Thr Leu Phe Trp Ala Ile Phe Gly Leu Ser Glu Val
        675                 680                 685

aaa tca gtg gtc atc aac tat aac cac aaa ttc att gaa aac att ggt     2112
Lys Ser Val Val Ile Asn Tyr Asn His Lys Phe Ile Glu Asn Ile Gly
    690                 695                 700

tac gtt ctt tat gga gtc tat aat gtt acg atg gtc att gtt ttg cta     2160
Tyr Val Leu Tyr Gly Val Tyr Asn Val Thr Met Val Ile Val Leu Leu
705                 710                 715                 720

aat atg tta att gcc atg atc aac agt tca ttc cag gaa att gag gat     2208
Asn Met Leu Ile Ala Met Ile Asn Ser Ser Phe Gln Glu Ile Glu Asp
                725                 730                 735

gac gct gat gtg gag tgg aaa ttt gca agg gcc aaa ctc tgg ttt tcc     2256
Asp Ala Asp Val Glu Trp Lys Phe Ala Arg Ala Lys Leu Trp Phe Ser
            740                 745                 750

tac ttt gag gag ggc aga aca ctt cct gta ccc ttc aat ctg gtg ccg     2304
Tyr Phe Glu Glu Gly Arg Thr Leu Pro Val Pro Phe Asn Leu Val Pro
        755                 760                 765

agt cca aag tcc ctg ttt tat ctc tta ctg aag ctt aaa aaa tgg att     2352
Ser Pro Lys Ser Leu Phe Tyr Leu Leu Leu Lys Leu Lys Lys Trp Ile
    770                 775                 780

tct gag ctg ttc cag ggc cat aaa aaa ggt ttc cag gaa gat gca gag     2400
Ser Glu Leu Phe Gln Gly His Lys Lys Gly Phe Gln Glu Asp Ala Glu
785                 790                 795                 800

atg aac aag ata aat gaa gaa aag aaa ctt gga att tta gga agt cat     2448
Met Asn Lys Ile Asn Glu Glu Lys Lys Leu Gly Ile Leu Gly Ser His
                805                 810                 815

gaa gac ctt tca aaa tta tca ctt gac aaa aaa cag gtt ggg cac aat     2496
Glu Asp Leu Ser Lys Leu Ser Leu Asp Lys Lys Gln Val Gly His Asn
            820                 825                 830

aaa caa cca agt ata agg agc tca gaa gat ttc cat cta aat agt ttc     2544
Lys Gln Pro Ser Ile Arg Ser Ser Glu Asp Phe His Leu Asn Ser Phe
        835                 840                 845

aat aat cct cca aga caa tat cag aaa ata atg aaa agg ctc att aaa     2592
Asn Asn Pro Pro Arg Gln Tyr Gln Lys Ile Met Lys Arg Leu Ile Lys
    850                 855                 860

aga tat gta ctg cag gcc cag ata gat aag gag agt gat gaa gtg aac     2640
```

-continued

```
Arg Tyr Val Leu Gln Ala Gln Ile Asp Lys Glu Ser Asp Glu Val Asn
865                 870                 875                 880

gaa ggg gaa ctg aag gaa att aag cag gac atc tca agt ctc cgc tat     2688
Glu Gly Glu Leu Lys Glu Ile Lys Gln Asp Ile Ser Ser Leu Arg Tyr
                885                 890                 895

gaa ctc ctt gaa gaa aaa tct cag aat aca gaa gac cta gca gaa ctt     2736
Glu Leu Leu Glu Glu Lys Ser Gln Asn Thr Glu Asp Leu Ala Glu Leu
            900                 905                 910

att aga gaa ctt gga gag aaa tta tcc atg gaa cca aat caa gag gaa     2784
Ile Arg Glu Leu Gly Glu Lys Leu Ser Met Glu Pro Asn Gln Glu Glu
        915                 920                 925

acc aat aga taa                                                     2796
Thr Asn Arg  *
    930


<210> SEQ ID NO 6
<211> LENGTH: 931
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 6

Met Ser Gln Ser Pro Ala Phe Gly Pro Arg Arg Gly Ile Ser Pro Arg
 1               5                  10                  15

Gly Ala Ala Gly Ala Ala Ala Arg Arg Asn Glu Ser Gln Asp Tyr Leu
            20                  25                  30

Leu Met Asp Ser Glu Leu Gly Glu Asp Gly Cys Pro Gln Ala Pro Leu
        35                  40                  45

Pro Cys Tyr Gly Tyr Tyr Pro Cys Phe Arg Gly Ser Asp Asn Arg Leu
    50                  55                  60

Ala His Arg Arg Gln Thr Val Leu Arg Glu Lys Gly Arg Arg Leu Ala
65                  70                  75                  80

Asn Arg Gly Pro Ala Tyr Met Phe Ser Asp Arg Ser Thr Ser Leu Ser
                85                  90                  95

Ile Glu Glu Glu Arg Phe Leu Asp Ala Ala Glu Tyr Gly Asn Ile Pro
            100                 105                 110

Val Val Arg Lys Met Leu Glu Glu Cys His Ser Leu Asn Val Asn Cys
        115                 120                 125

Val Asp Tyr Met Gly Gln Asn Ala Leu Gln Leu Ala Val Ala Asn Glu
    130                 135                 140

His Leu Glu Ile Thr Glu Leu Leu Leu Lys Lys Glu Asn Leu Ser Arg
145                 150                 155                 160

Ile Gly Asp Ala Leu Leu Leu Ala Ile Ser Lys Gly Tyr Val Arg Ile
                165                 170                 175

Val Glu Ala Ile Leu Ser His Pro Ala Phe Ala Glu Gly Lys Arg Leu
            180                 185                 190

Ala Thr Ser Pro Ser Gln Ser Glu Leu Gln Gln Asp Asp Phe Tyr Ala
        195                 200                 205

Tyr Asp Glu Asp Gly Thr Arg Ser Ser His Asp Val Thr Pro Ile Ile
    210                 215                 220

Leu Ala Ala His Cys Arg Glu Tyr Glu Ile Val His Thr Leu Leu Arg
225                 230                 235                 240

Lys Gly Ala Arg Ile Glu Arg Pro His Asp Tyr Phe Cys Lys Cys Asn
                245                 250                 255

Asp Cys Asn Gln Lys Gln Lys His Asp Ser Phe Ser His Ser Arg Ser
```

-continued

```
            260                 265                 270

Arg Ile Asn Ala Tyr Lys Gly Leu Ala Ser Pro Ala Tyr Leu Ser Leu
        275                 280                 285

Ser Ser Glu Asp Pro Val Met Thr Ala Leu Glu Leu Ser Asn Glu Leu
    290                 295                 300

Ala Val Leu Ala Asn Ile Glu Lys Glu Phe Lys Asn Asp Tyr Lys Lys
305                 310                 315                 320

Leu Ser Met Gln Cys Lys Asp Phe Val Val Gly Leu Leu Asp Leu Cys
                325                 330                 335

Arg Asn Thr Glu Glu Val Glu Ala Ile Leu Asn Gly Asp Val Glu Thr
            340                 345                 350

Leu Gln Ser Gly Asp His Gly Arg Pro Asn Leu Ser Arg Leu Lys Leu
        355                 360                 365

Ala Ile Lys Tyr Glu Val Lys Lys Phe Val Ala His Pro Asn Cys Gln
    370                 375                 380

Gln Gln Leu Leu Ser Ile Trp Tyr Glu Asn Leu Ser Gly Leu Arg Gln
385                 390                 395                 400

Gln Thr Met Ala Val Lys Phe Leu Val Ala Leu Ala Val Ala Ile Gly
                405                 410                 415

Leu Pro Phe Leu Ala Leu Ile Tyr Trp Phe Ala Pro Cys Ser Lys Met
            420                 425                 430

Gly Lys Ile Met Arg Gly Pro Phe Met Lys Phe Val Ala His Ala Ala
        435                 440                 445

Ser Phe Thr Ile Phe Leu Gly Leu Leu Val Met Asn Ala Ala Asp Arg
    450                 455                 460

Phe Glu Gly Thr Lys Leu Leu Pro Asn Glu Thr Ser Thr Asp Asn Ala
465                 470                 475                 480

Lys Gln Leu Phe Arg Met Lys Thr Ser Cys Phe Ser Trp Met Glu Met
                485                 490                 495

Leu Ile Ile Ser Trp Val Ile Gly Met Ile Trp Ala Glu Cys Lys Glu
            500                 505                 510

Ile Trp Thr Gln Gly Pro Lys Glu Tyr Leu Phe Glu Leu Trp Asn Met
        515                 520                 525

Leu Asp Phe Gly Met Leu Ala Ile Phe Pro Ala Ser Phe Ile Ala Arg
    530                 535                 540

Phe Met Ala Phe Trp His Ala Ser Lys Ala Gln Ser Ile Ile Asp Ala
545                 550                 555                 560

Asn Asp Thr Leu Lys Asp Leu Thr Lys Val Thr Leu Gly Asp Asn Val
                565                 570                 575

Lys Tyr Tyr Asn Leu Ala Arg Ile Lys Trp Asp Pro Ser Asp Pro Gln
            580                 585                 590

Ile Ile Ser Glu Gly Leu Tyr Ala Ile Ala Val Val Leu Ser Phe Ser
        595                 600                 605

Arg Ile Ala Tyr Ile Leu Pro Ala Asn Glu Ser Phe Gly Pro Leu Gln
    610                 615                 620

Ile Ser Leu Gly Arg Thr Val Lys Asp Ile Phe Lys Phe Met Val Ile
625                 630                 635                 640

Phe Ile Met Val Phe Val Ala Phe Met Ile Gly Met Phe Asn Leu Tyr
                645                 650                 655

Ser Tyr Tyr Ile Gly Ala Lys Gln Asn Glu Ala Phe Thr Thr Val Glu
            660                 665                 670
```

-continued

```
Glu Ser Phe Lys Thr Leu Phe Trp Ala Ile Phe Gly Leu Ser Glu Val
        675                 680                 685

Lys Ser Val Val Ile Asn Tyr Asn His Lys Phe Ile Glu Asn Ile Gly
    690                 695                 700

Tyr Val Leu Tyr Gly Val Tyr Asn Val Thr Met Val Ile Val Leu Leu
705                 710                 715                 720

Asn Met Leu Ile Ala Met Ile Asn Ser Ser Phe Gln Glu Ile Glu Asp
                725                 730                 735

Asp Ala Asp Val Glu Trp Lys Phe Ala Arg Ala Lys Leu Trp Phe Ser
            740                 745                 750

Tyr Phe Glu Glu Gly Arg Thr Leu Pro Val Pro Phe Asn Leu Val Pro
        755                 760                 765

Ser Pro Lys Ser Leu Phe Tyr Leu Leu Leu Lys Leu Lys Lys Trp Ile
    770                 775                 780

Ser Glu Leu Phe Gln Gly His Lys Lys Gly Phe Gln Glu Asp Ala Glu
785                 790                 795                 800

Met Asn Lys Ile Asn Glu Glu Lys Lys Leu Gly Ile Leu Gly Ser His
                805                 810                 815

Glu Asp Leu Ser Lys Leu Ser Leu Asp Lys Lys Gln Val Gly His Asn
            820                 825                 830

Lys Gln Pro Ser Ile Arg Ser Ser Glu Asp Phe His Leu Asn Ser Phe
        835                 840                 845

Asn Asn Pro Pro Arg Gln Tyr Gln Lys Ile Met Lys Arg Leu Ile Lys
    850                 855                 860

Arg Tyr Val Leu Gln Ala Gln Ile Asp Lys Glu Ser Asp Glu Val Asn
865                 870                 875                 880

Glu Gly Glu Leu Lys Glu Ile Lys Gln Asp Ile Ser Ser Leu Arg Tyr
                885                 890                 895

Glu Leu Leu Glu Glu Lys Ser Gln Asn Thr Glu Asp Leu Ala Glu Leu
            900                 905                 910

Ile Arg Glu Leu Gly Glu Lys Leu Ser Met Glu Pro Asn Gln Glu Glu
        915                 920                 925

Thr Asn Arg
    930


<210> SEQ ID NO 7
<211> LENGTH: 2796
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(2796)

<400> SEQUENCE: 7

atg agc caa agc ccg gcg ttc ggg ccc cgg agg ggc att tct ccc cgg       48
Met Ser Gln Ser Pro Ala Phe Gly Pro Arg Arg Gly Ile Ser Pro Arg
 1               5                  10                  15

ggc gct gcc gga gcc gct gcg cgg cgc aac gag agc cag gac tat ctg       96
Gly Ala Ala Gly Ala Ala Ala Arg Arg Asn Glu Ser Gln Asp Tyr Leu
            20                  25                  30

ctc atg gac tcg gag ctg gga gaa gac ggc tgc ccg caa gcc ccg ctg      144
Leu Met Asp Ser Glu Leu Gly Glu Asp Gly Cys Pro Gln Ala Pro Leu
        35                  40                  45

cct tgc tac ggc tac tac ccc tgc ttc cgg gga tct gac aac aga ctg      192
```

-continued

```
Pro Cys Tyr Gly Tyr Tyr Pro Cys Phe Arg Gly Ser Asp Asn Arg Leu
     50                  55                  60

gct cac cgg cgg cag aca gtt ctc cgt gag aag ggg aga agg tta gct      240
Ala His Arg Arg Gln Thr Val Leu Arg Glu Lys Gly Arg Arg Leu Ala
 65                  70                  75                  80

aat cga gga cca gca tac atg ttt agt gat cgc tcc aca agc cta tct      288
Asn Arg Gly Pro Ala Tyr Met Phe Ser Asp Arg Ser Thr Ser Leu Ser
                 85                  90                  95

ata gag gag gaa cgc ttt ttg gat gca gct gaa tat ggt aac atc cca      336
Ile Glu Glu Glu Arg Phe Leu Asp Ala Ala Glu Tyr Gly Asn Ile Pro
            100                 105                 110

gtg gtg cgg aag atg tta gaa gaa tgc cac tca ctc aac gtt aac tgt      384
Val Val Arg Lys Met Leu Glu Glu Cys His Ser Leu Asn Val Asn Cys
        115                 120                 125

gtg gat tac atg ggc cag aat gcc cta cag ttg gca gtg gcc aat gag      432
Val Asp Tyr Met Gly Gln Asn Ala Leu Gln Leu Ala Val Ala Asn Glu
    130                 135                 140

cat ctg gaa att aca gaa ctt ctt ctc aag aaa gaa aac ctc tct cga      480
His Leu Glu Ile Thr Glu Leu Leu Leu Lys Lys Glu Asn Leu Ser Arg
145                 150                 155                 160

att ggg gat gct ttg ctt cta gct att agt aaa ggt tat gtt cgg att      528
Ile Gly Asp Ala Leu Leu Leu Ala Ile Ser Lys Gly Tyr Val Arg Ile
                165                 170                 175

gtg gaa gca att ctc agt cat ccg gct ttt gct gaa ggc aag agg tta      576
Val Glu Ala Ile Leu Ser His Pro Ala Phe Ala Glu Gly Lys Arg Leu
            180                 185                 190

gca acc agc cct agc cag tct gaa ctc cag caa gat gat ttt tat gcc      624
Ala Thr Ser Pro Ser Gln Ser Glu Leu Gln Gln Asp Asp Phe Tyr Ala
        195                 200                 205

tat gat gaa gat ggg aca cgg tcc tcc cat gat gtg act cca atc att      672
Tyr Asp Glu Asp Gly Thr Arg Ser Ser His Asp Val Thr Pro Ile Ile
    210                 215                 220

ctg gct gcc cac tgc cgg gaa tat gaa att gtg cat acc ctc ctg cgg      720
Leu Ala Ala His Cys Arg Glu Tyr Glu Ile Val His Thr Leu Leu Arg
225                 230                 235                 240

aag ggt gct agg att gaa cgg cct cat gat tat ttc tgc aag tgc aat      768
Lys Gly Ala Arg Ile Glu Arg Pro His Asp Tyr Phe Cys Lys Cys Asn
                245                 250                 255

gac tgc aac cag aaa cag aag cat gac tcg ttt agc cac tcc aga tct      816
Asp Cys Asn Gln Lys Gln Lys His Asp Ser Phe Ser His Ser Arg Ser
            260                 265                 270

agg att aat gcc tat aaa ggc ctg gca agt ccg gct tac ctg tca ttg      864
Arg Ile Asn Ala Tyr Lys Gly Leu Ala Ser Pro Ala Tyr Leu Ser Leu
        275                 280                 285

tct agt gaa gat cca gtc atg acg gct tta gaa ctt agc aat gaa ctg      912
Ser Ser Glu Asp Pro Val Met Thr Ala Leu Glu Leu Ser Asn Glu Leu
    290                 295                 300

gca gtt ctg gcc aat att gag aaa gag ttc aag aat gac tac aaa aaa      960
Ala Val Leu Ala Asn Ile Glu Lys Glu Phe Lys Asn Asp Tyr Lys Lys
305                 310                 315                 320

ctg tca atg cag tgc aaa gac ttt gtt gtt gga ctc ctt gat ctg tgc     1008
Leu Ser Met Gln Cys Lys Asp Phe Val Val Gly Leu Leu Asp Leu Cys
                325                 330                 335

aga aac act gaa gaa gtc gag gcc att ctg aat ggg gat gtt gaa acg     1056
Arg Asn Thr Glu Glu Val Glu Ala Ile Leu Asn Gly Asp Val Glu Thr
            340                 345                 350

ctc cag agt ggt gat cac ggt cgc cca aat ctc agc cgt tta aaa ctt     1104
```

-continued

```
Leu Gln Ser Gly Asp His Gly Arg Pro Asn Leu Ser Arg Leu Lys Leu
        355                 360                 365

gcc att aaa tat gaa gta aaa aaa ttt gta gct cat cca aac tgc caa     1152
Ala Ile Lys Tyr Glu Val Lys Lys Phe Val Ala His Pro Asn Cys Gln
    370                 375                 380

cag caa ctt ctc tcc att tgg tat gag aat ctt tct ggt tta cga cag     1200
Gln Gln Leu Leu Ser Ile Trp Tyr Glu Asn Leu Ser Gly Leu Arg Gln
385                 390                 395                 400

cag aca atg gcg gtc aag ttc ctt gtg gcc ctt gct gtt gcc att gga     1248
Gln Thr Met Ala Val Lys Phe Leu Val Ala Leu Ala Val Ala Ile Gly
                405                 410                 415

ctg ccc ttc ctg gct ctc att tac tgg ttt gct cca tgc agc aag atg     1296
Leu Pro Phe Leu Ala Leu Ile Tyr Trp Phe Ala Pro Cys Ser Lys Met
            420                 425                 430

ggg aag ata atg cgt gga cca ttc atg aag ttt gta gca cac gca gcc     1344
Gly Lys Ile Met Arg Gly Pro Phe Met Lys Phe Val Ala His Ala Ala
        435                 440                 445

tcc ttc acc att ttt ctg gga ctg cta gtc atg aat gca gct gac aga     1392
Ser Phe Thr Ile Phe Leu Gly Leu Leu Val Met Asn Ala Ala Asp Arg
    450                 455                 460

ttt gaa ggc aca aaa ctc ctt cct aat gaa acc agc aca gat aat gca     1440
Phe Glu Gly Thr Lys Leu Leu Pro Asn Glu Thr Ser Thr Asp Asn Ala
465                 470                 475                 480

aaa cag ctg ttc agg atg aaa aca tcc tgc ttc tca tgg atg gag atg     1488
Lys Gln Leu Phe Arg Met Lys Thr Ser Cys Phe Ser Trp Met Glu Met
                485                 490                 495

ctc att ata tcc tgg gta ata ggc atg ata tgg gct gaa tgt aaa gaa     1536
Leu Ile Ile Ser Trp Val Ile Gly Met Ile Trp Ala Glu Cys Lys Glu
            500                 505                 510

atc tgg act cag ggc ccc aag gaa tat ttg ttt gag ttg tgg aac atg     1584
Ile Trp Thr Gln Gly Pro Lys Glu Tyr Leu Phe Glu Leu Trp Asn Met
        515                 520                 525

ctt gat ttt ggt atg tta gca att ttc cca gca tca ttc att gcg aga     1632
Leu Asp Phe Gly Met Leu Ala Ile Phe Pro Ala Ser Phe Ile Ala Arg
    530                 535                 540

ttc atg gca ttt tgg cat gct tcc aaa gcc cag agc atc att gac gca     1680
Phe Met Ala Phe Trp His Ala Ser Lys Ala Gln Ser Ile Ile Asp Ala
545                 550                 555                 560

aac gat act ttg aag gac ttg acg aaa gta aca ttg gga gac aat gtg     1728
Asn Asp Thr Leu Lys Asp Leu Thr Lys Val Thr Leu Gly Asp Asn Val
                565                 570                 575

aaa tac tac aat ttg gcc agg ata aag tgg gac ccc tct gat cct caa     1776
Lys Tyr Tyr Asn Leu Ala Arg Ile Lys Trp Asp Pro Ser Asp Pro Gln
            580                 585                 590

ata ata tct gaa ggt ctt tat gca att gct gta gtt tta agt ttc tct     1824
Ile Ile Ser Glu Gly Leu Tyr Ala Ile Ala Val Val Leu Ser Phe Ser
        595                 600                 605

agg ata gct tat att tta cca gca aat gaa agc ttt gga cct ctg cag     1872
Arg Ile Ala Tyr Ile Leu Pro Ala Asn Glu Ser Phe Gly Pro Leu Gln
    610                 615                 620

ata tca ctt gga aga aca gtc aaa gac atc ttc aag ttc atg gtc ata     1920
Ile Ser Leu Gly Arg Thr Val Lys Asp Ile Phe Lys Phe Met Val Ile
625                 630                 635                 640

ttc att atg gtg ttt gtg gcc ttt atg att gga atg ttc aat ctc tac     1968
Phe Ile Met Val Phe Val Ala Phe Met Ile Gly Met Phe Asn Leu Tyr
                645                 650                 655

tcc tac tac att ggt gca aaa caa aat gaa gcc ttc aca aca gtt gaa     2016
```

-continued

```
Ser Tyr Tyr Ile Gly Ala Lys Gln Asn Glu Ala Phe Thr Thr Val Glu
            660                 665                 670

gag agt ttt aag aca ctg ttc tgg gct ata ttt gga ctt tct gaa gtg     2064
Glu Ser Phe Lys Thr Leu Phe Trp Ala Ile Phe Gly Leu Ser Glu Val
        675                 680                 685

aaa tca gtg gtc atc aac tat aac cac aaa ttc att gaa aac att ggt     2112
Lys Ser Val Val Ile Asn Tyr Asn His Lys Phe Ile Glu Asn Ile Gly
    690                 695                 700

tac gtt ctt tat gga gtc tat aat gtt acg atg gtc att gtt ttg cta     2160
Tyr Val Leu Tyr Gly Val Tyr Asn Val Thr Met Val Ile Val Leu Leu
705                 710                 715                 720

aat atg tta att gcc atg atc aac agt tca ttc cag gaa att gag gat     2208
Asn Met Leu Ile Ala Met Ile Asn Ser Ser Phe Gln Glu Ile Glu Asp
                725                 730                 735

gac gct gat gtg gag tgg aaa ttt gca agg gcc aaa ctc tgg ttt tcc     2256
Asp Ala Asp Val Glu Trp Lys Phe Ala Arg Ala Lys Leu Trp Phe Ser
            740                 745                 750

tac ttt gag gag ggc aga aca ctt cct gta ccc ttc aat ctg gtg ccg     2304
Tyr Phe Glu Glu Gly Arg Thr Leu Pro Val Pro Phe Asn Leu Val Pro
        755                 760                 765

agt cca aag tcc ctg ttt tat ctc tta ctg aag ctt aaa aaa tgg att     2352
Ser Pro Lys Ser Leu Phe Tyr Leu Leu Leu Lys Leu Lys Lys Trp Ile
    770                 775                 780

tct gag ctg ttc cag ggc cat aaa aaa ggt ttc cag gaa gat gca gag     2400
Ser Glu Leu Phe Gln Gly His Lys Lys Gly Phe Gln Glu Asp Ala Glu
785                 790                 795                 800

atg aac aag ata aat gaa gaa aag aaa ctt gga att tta gga agt cat     2448
Met Asn Lys Ile Asn Glu Glu Lys Lys Leu Gly Ile Leu Gly Ser His
                805                 810                 815

gaa gac ctt tca aaa tta tca ctt gac aaa aaa cag gtt ggg cac aat     2496
Glu Asp Leu Ser Lys Leu Ser Leu Asp Lys Lys Gln Val Gly His Asn
            820                 825                 830

aaa caa cca agt ata agg agc tca gaa gat ttc cat cta aat agt ttc     2544
Lys Gln Pro Ser Ile Arg Ser Ser Glu Asp Phe His Leu Asn Ser Phe
        835                 840                 845

aat aat cct cca aga caa tat cag aaa ata atg aaa agg ctc att aaa     2592
Asn Asn Pro Pro Arg Gln Tyr Gln Lys Ile Met Lys Arg Leu Ile Lys
    850                 855                 860

aga tat gta ctg cag gcc cag ata gat aag gag agt gat gaa gtg aac     2640
Arg Tyr Val Leu Gln Ala Gln Ile Asp Lys Glu Ser Asp Glu Val Asn
865                 870                 875                 880

gaa ggg gaa ctg aag gaa att aag cag gac atc tca agt ctc cgc tat     2688
Glu Gly Glu Leu Lys Glu Ile Lys Gln Asp Ile Ser Ser Leu Arg Tyr
                885                 890                 895

gaa ctc ctt gaa gaa aaa tct cag aat aca gaa gac cta gca gaa ctt     2736
Glu Leu Leu Glu Glu Lys Ser Gln Asn Thr Glu Asp Leu Ala Glu Leu
            900                 905                 910

att aga gaa ctt gga gag aaa tta tcc atg gaa cca aat caa gag gaa     2784
Ile Arg Glu Leu Gly Glu Lys Leu Ser Met Glu Pro Asn Gln Glu Glu
        915                 920                 925

acc aat aga taa                                                     2796
Thr Asn Arg  *
    930
```

<210> SEQ ID NO 8
<211> LENGTH: 35
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence

-continued

<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (24)...(25)
<223> OTHER INFORMATION: Xaa = Any Amino Acid

<400> SEQUENCE: 8

Asp Gly Arg Thr Pro Leu His Leu Ala Ala Arg Asn Gly His Leu Glu
1 5 10 15

Val Val Lys Leu Leu Leu Glu Xaa Xaa Ala Gly Ala Asp Val Asn Ala
20 25 30

Arg Asp Lys
35

<210> SEQ ID NO 9
<211> LENGTH: 33
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 9

Asp Gly Arg Thr Pro Leu His Leu Ala Ala Arg Asn Gly His Leu Glu
1 5 10 15

Val Val Lys Leu Leu Leu Glu Ala Gly Ala Asp Val Asn Ala Arg Asp
20 25 30

Lys

<210> SEQ ID NO 10
<211> LENGTH: 33
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 10

Asp Gly Arg Thr Pro Leu His Leu Ala Ala Arg Asn Gly His Leu Glu
1 5 10 15

Val Val Lys Leu Leu Leu Glu Ala Gly Ala Asp Val Asn Ala Arg Asp
20 25 30

Lys

<210> SEQ ID NO 11
<211> LENGTH: 250
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (67)...(67)
<223> OTHER INFORMATION: Xaa = Any Amino Acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (71)...(92)
<223> OTHER INFORMATION: Xaa = Any Amino Acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (216)...(224)
<223> OTHER INFORMATION: Xaa = Any Amino Acid

<400> SEQUENCE: 11

-continued

```
Ile Leu Phe Ile Leu Asp Leu Leu Phe Val Leu Leu Phe Leu Leu Glu
 1               5                  10                  15

Ile Val Leu Lys Phe Ile Ala Tyr Gly Leu Lys Ser Thr Ser Asn Ile
            20                  25                  30

Ala Ala Lys Tyr Leu Lys Ser Ile Phe Asn Ile Leu Asp Leu Leu Ala
        35                  40                  45

Ile Leu Pro Leu Leu Leu Leu Leu Val Leu Phe Leu Ser Gly Thr Glu
    50                  55                  60

Gln Val Xaa Lys Arg Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
65                  70                  75                  80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Glu Arg Ser
                85                  90                  95

Leu Glu Leu Ser Gln Tyr Arg Ile Leu Arg Phe Leu Arg Leu Leu Arg
            100                 105                 110

Leu Leu Arg Leu Leu Arg Leu Leu Arg Leu Leu Arg Arg Leu Glu Thr
        115                 120                 125

Leu Phe Glu Phe Glu Leu Gly Thr Leu Ala Trp Ser Leu Gln Ser Leu
    130                 135                 140

Gly Arg Ala Leu Lys Ser Ile Leu Arg Phe Leu Leu Leu Leu Leu Leu
145                 150                 155                 160

Leu Leu Ile Gly Phe Ser Val Ile Gly Tyr Leu Leu Phe Lys Gly Tyr
                165                 170                 175

Glu Asp Leu Ser Glu Asn Glu Val Asp Gly Asn Ser Glu Phe Ser Ser
            180                 185                 190

Tyr Phe Asp Ala Phe Tyr Phe Leu Phe Val Thr Leu Thr Thr Val Gly
        195                 200                 205

Phe Gly Asp Leu Val Pro Val Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
    210                 215                 220

Trp Leu Gly Ile Ile Phe Phe Val Leu Phe Phe Ile Ile Val Gly Leu
225                 230                 235                 240

Leu Leu Leu Asn Leu Leu Ile Ala Val Ile
                245                 250


<210> SEQ ID NO 12
<211> LENGTH: 227
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 12

Asn Val Asp Glu Ser Cys Glu Asp Arg Ser Trp Arg Ser Ala Leu Leu
 1               5                  10                  15

Ile Ala Ile Glu Asn Glu Asn Leu Glu Ile Met Glu Leu Leu Leu Asn
            20                  25                  30

His Ser Val Phe Asp Phe Arg Arg Val Gly Asp Ala Leu Leu Tyr Ala
        35                  40                  45

Ile Ser Lys Glu Tyr Val Arg Ala Val Glu Ala Leu Leu Ser His Pro
    50                  55                  60

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

Leu Gln Gln Asp Asp Phe Tyr Ala Tyr Asp Glu Asp Gly Ser Glu Phe
                85                  90                  95
```

-continued

```
Thr Pro Asp Ile Thr Pro Ile Ile Leu Ala Ala His Thr Asn Asn Tyr
            100                 105                 110

Glu Ile Ile Lys Leu Leu Leu Gln Lys Gly Val Ser Ile Pro Arg Pro
        115                 120                 125

His Asp Val Arg Cys Asn Cys Thr Glu Cys Val Ser Lys Gln Glu Val
    130                 135                 140

Asp Ser Leu Arg His Ser Arg Ser Arg Leu Asn Ile Tyr Lys Ala Leu
145                 150                 155                 160

Ala Ser Pro Ser Leu Ile Ala Leu Ser Ser Glu Asp Pro Ile Leu Thr
                165                 170                 175

Ala Phe Glu Leu Ser Trp Glu Leu Lys Glu Leu Ser Lys Val Glu Asn
            180                 185                 190

Glu Phe Lys Asn Glu Tyr Glu Glu Leu Ser Arg Gln Cys Lys Gln Phe
        195                 200                 205

Ala Lys Asp Leu Leu Asp Gln Cys Arg Asn Ser Arg Glu Leu Glu Val
    210                 215                 220

Ile Leu Asn
225


<210> SEQ ID NO 13
<211> LENGTH: 2875
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (115)...(1869)

<400> SEQUENCE: 13

atgttgtccc cctttctcac tcccaccttg caggccaccg gagccgccag ctgttggaac     60

tgagctactg cagaaaggga agtggagagt aagggccagg ccccgtgggg gcag atg      117
Met
 1

gcc ggc aga agg ctg aat ctg cgc tgg gca ctg agt gtg ctt tgt gtg      165
Ala Gly Arg Arg Leu Asn Leu Arg Trp Ala Leu Ser Val Leu Cys Val
             5                  10                  15

ctg cta atg gcg gag aca gtg tct ggg act agg ggc tcg tct aca gga      213
Leu Leu Met Ala Glu Thr Val Ser Gly Thr Arg Gly Ser Ser Thr Gly
         20                  25                  30

gct cac att agc ccc cag ttt cca gct tca ggt gtg aac cag acc ccc      261
Ala His Ile Ser Pro Gln Phe Pro Ala Ser Gly Val Asn Gln Thr Pro
     35                  40                  45

gtg gta gac tgc cgc aag gtg tgt ggc ctg aat gtc tct gac cgc tgt      309
Val Val Asp Cys Arg Lys Val Cys Gly Leu Asn Val Ser Asp Arg Cys
 50                  55                  60                  65

gac ttc atc cgg acc aac cct gac tgc cac agt gat ggg ggg tac ctg      357
Asp Phe Ile Arg Thr Asn Pro Asp Cys His Ser Asp Gly Gly Tyr Leu
                 70                  75                  80

gac tac ctg gaa ggc atc ttc tgc cac ttc cct ccc agc ctc ctc cct      405
Asp Tyr Leu Glu Gly Ile Phe Cys His Phe Pro Pro Ser Leu Leu Pro
             85                  90                  95

ctg gct gtc act ctc tac gtt tcc tgg ctg ctc tac ctg ttt ctg att      453
Leu Ala Val Thr Leu Tyr Val Ser Trp Leu Leu Tyr Leu Phe Leu Ile
        100                 105                 110

ctg gga gtc acc gca gcc aag ttt ttc tgc ccc aac ttg tcg gcc att      501
Leu Gly Val Thr Ala Ala Lys Phe Phe Cys Pro Asn Leu Ser Ala Ile
    115                 120                 125
```

-continued

```
tct acc aca ctg aag ctc tcc cac aac gtg gca ggc gtc acc ttc ctg      549
Ser Thr Thr Leu Lys Leu Ser His Asn Val Ala Gly Val Thr Phe Leu
130                 135                 140                 145

gca ttt ggg aat ggt gca cct gac atc ttc agt gcc ctg gtg gcc ttc      597
Ala Phe Gly Asn Gly Ala Pro Asp Ile Phe Ser Ala Leu Val Ala Phe
                150                 155                 160

tct gac ccg cac aca gcc ggc ctg gcc ctt ggg gca ctg ttt ggc gct      645
Ser Asp Pro His Thr Ala Gly Leu Ala Leu Gly Ala Leu Phe Gly Ala
            165                 170                 175

ggc gtg ctg gtt acc aca gtg gtg gcc gga ggc att acc atc cta cac      693
Gly Val Leu Val Thr Thr Val Val Ala Gly Gly Ile Thr Ile Leu His
        180                 185                 190

ccc ttc atg gct gcc tcc agg ccc ttc ttc agg gac atc gtt ttc tac      741
Pro Phe Met Ala Ala Ser Arg Pro Phe Phe Arg Asp Ile Val Phe Tyr
    195                 200                 205

atg gtg gct gtg ttc ctg acc ttc ctc atg ctc ttc cgt ggc agg gtc      789
Met Val Ala Val Phe Leu Thr Phe Leu Met Leu Phe Arg Gly Arg Val
210                 215                 220                 225

acc ctg gca tgg gct ctg ggt tac ctg ggc ttg tat gtg ttc tat gtg      837
Thr Leu Ala Trp Ala Leu Gly Tyr Leu Gly Leu Tyr Val Phe Tyr Val
                230                 235                 240

gtc act gtg att ctc tgc acc tgg atc tac caa cgg caa cgg aga gga      885
Val Thr Val Ile Leu Cys Thr Trp Ile Tyr Gln Arg Gln Arg Arg Gly
            245                 250                 255

tct ctg ttc tgc ccc atg cca gtt act cca gag atc ctc tca gac tcc      933
Ser Leu Phe Cys Pro Met Pro Val Thr Pro Glu Ile Leu Ser Asp Ser
        260                 265                 270

gag gag gac cgg gta tct tct aat acc aac agc tat gac tac ggt gat      981
Glu Glu Asp Arg Val Ser Ser Asn Thr Asn Ser Tyr Asp Tyr Gly Asp
    275                 280                 285

gag tac cgg ccg ctg ttc ttc tac cag gag acc acg gct cag atc ctg     1029
Glu Tyr Arg Pro Leu Phe Phe Tyr Gln Glu Thr Thr Ala Gln Ile Leu
290                 295                 300                 305

gtc cgg gcc ctc aat ccc ctg gat tac atg aag tgg aga agg aaa tca     1077
Val Arg Ala Leu Asn Pro Leu Asp Tyr Met Lys Trp Arg Arg Lys Ser
                310                 315                 320

gca tac tgg aaa gcc ctc aag gtg ttc aag ctg cct gtg gag ttc ctg     1125
Ala Tyr Trp Lys Ala Leu Lys Val Phe Lys Leu Pro Val Glu Phe Leu
            325                 330                 335

ctg ctc ctc aca gtc ccc gtc gtg gac ccg gac aag gat gac cag aac     1173
Leu Leu Leu Thr Val Pro Val Val Asp Pro Asp Lys Asp Asp Gln Asn
        340                 345                 350

tgg aaa cgg ccc ctc aac tgt ctg cat ctg gtt atc agc ccc ctg gtt     1221
Trp Lys Arg Pro Leu Asn Cys Leu His Leu Val Ile Ser Pro Leu Val
    355                 360                 365

gtg gtc ctg acc ctg cag tcg ggg acc tat ggt gtc tat gag ata ggc     1269
Val Val Leu Thr Leu Gln Ser Gly Thr Tyr Gly Val Tyr Glu Ile Gly
370                 375                 380                 385

ggc ctc gtt ccc gtc tgg gtc gtg gtg gtg atc gca ggc aca gcc ttg     1317
Gly Leu Val Pro Val Trp Val Val Val Val Ile Ala Gly Thr Ala Leu
                390                 395                 400

gct tca gtg acc ttt ttt gcc aca tct gac agc cag ccc ccc agg ctt     1365
Ala Ser Val Thr Phe Phe Ala Thr Ser Asp Ser Gln Pro Pro Arg Leu
            405                 410                 415

cac tgg ctc ttt gct ttc ctg ggc ttt ctg acc agc gcc ctg tgg atc     1413
His Trp Leu Phe Ala Phe Leu Gly Phe Leu Thr Ser Ala Leu Trp Ile
        420                 425                 430
```

-continued

```
aac gcg gcc gcc aca gag gtg gtg aac atc ttg cgg tcc ctg ggt gtg     1461
Asn Ala Ala Ala Thr Glu Val Val Asn Ile Leu Arg Ser Leu Gly Val
    435                 440                 445

gtc ttc cgg ctg agc aac act gtg ctg ggg ctc acg ctg ctg gcc tgg     1509
Val Phe Arg Leu Ser Asn Thr Val Leu Gly Leu Thr Leu Leu Ala Trp
450                 455                 460                 465

ggg aac agc att gga gat gcc ttc tcg gat ttc aca ctg gct cgc cag     1557
Gly Asn Ser Ile Gly Asp Ala Phe Ser Asp Phe Thr Leu Ala Arg Gln
                470                 475                 480

ggc tac cca cgg atg gcg ttc tcc gcc tgc ttt ggc ggc atc atc ttc     1605
Gly Tyr Pro Arg Met Ala Phe Ser Ala Cys Phe Gly Gly Ile Ile Phe
            485                 490                 495

aac atc ctc gtg ggt gtg ggg ctg ggc tgc ctg ctc cag atc tcc cga     1653
Asn Ile Leu Val Gly Val Gly Leu Gly Cys Leu Leu Gln Ile Ser Arg
        500                 505                 510

agc cac aca gaa gtg aag ctg gag cca gac gga ctg ctg gtg tgg gtc     1701
Ser His Thr Glu Val Lys Leu Glu Pro Asp Gly Leu Leu Val Trp Val
    515                 520                 525

ctg gca ggc gcc ctg ggg ctc agc ctc gtc ttc tcc ctg gtc tca gtc     1749
Leu Ala Gly Ala Leu Gly Leu Ser Leu Val Phe Ser Leu Val Ser Val
530                 535                 540                 545

cca ttg cag tgc ttc cag ctc agc aga gtc tat ggc ttc tgc ctg ctc     1797
Pro Leu Gln Cys Phe Gln Leu Ser Arg Val Tyr Gly Phe Cys Leu Leu
                550                 555                 560

ctc ttc tac ctg aac ttc ctt gtc gtg gcc ctc ctc act gaa ttt gga     1845
Leu Phe Tyr Leu Asn Phe Leu Val Val Ala Leu Leu Thr Glu Phe Gly
            565                 570                 575

gtg att cac ctg aaa agc atg tga ctgaagccgc ttagtgctgt ggcctcactg    1899
Val Ile His Leu Lys Ser Met  *
        580

caggcaggag ccccgcccct cctgccgggg gaggcccagg gaccggagca tttctgcaag   1959

gcccttgtgg gcacgagagt gcggcccttg ctgctggaga tctgaggtca ctgctgtgag   2019

ctgggagaac tgctgtgtac ctcttgctgc cagcacccaa cagccttgcc gtggggacct   2079

tggaaacctg gctttgctct ggacaaaggg ttccagagag aagctagaag tccccccttga  2139

atgaccccca gagcccctct gagaagggct ggagtttggg ggaaggggat ggctggatgt   2199

gctccaggcc atgctggagg taccccccgag gcacaggcac tgcccgttcc ccttgcctgg  2259

gcttcaggcc ttctggcacc ttctcaggac acaagtggct gcccaaccct gactcagaga   2319

atgagggtgg cttggacccc tgggaatcag gccgccgagg gctgagctcc agagccgcac   2379

catctgccac aaacagaatt cgagacatac ttaattttga atttctcctt gccacgttaa   2439

taaagccaaa agcagcgggt gctattcgtg gcaacacact tcactgaacc cacttgcttc   2499

caaaacgatg ccagcccgag gcactgctac gccagcagct gccacatggg atggtggctc   2559

aggcgctccc tccaggattc tgcccctgcc tgtccacaga ctcctttgtg ctggaacctg   2619

ggctcctcca gctgccaggc aggagtcggt aggactgtgc ctgtgcctcc ctcagcgggg   2679

ccctgggcgg ggttccaagg cctgcgagct gggaaaggac agatgagggg acctcgtgcc   2739

ttcttgctgt catgcaatga ccccgcctta tgttgccgaa ataagcaact cttaggtttg   2799

cctgactgcc ttatgctggt aaagaaaagg gattcaactg tctcttttcc aaataaaaaa   2859

aaaagtcaaa atttca                                                   2875
```

<210> SEQ ID NO 14

-continued

<211> LENGTH: 584
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 14

Met Ala Gly Arg Arg Leu Asn Leu Arg Trp Ala Leu Ser Val Leu Cys
1 5 10 15

Val Leu Leu Met Ala Glu Thr Val Ser Gly Thr Arg Gly Ser Ser Thr
20 25 30

Gly Ala His Ile Ser Pro Gln Phe Pro Ala Ser Gly Val Asn Gln Thr
35 40 45

Pro Val Val Asp Cys Arg Lys Val Cys Gly Leu Asn Val Ser Asp Arg
50 55 60

Cys Asp Phe Ile Arg Thr Asn Pro Asp Cys His Ser Asp Gly Gly Tyr
65 70 75 80

Leu Asp Tyr Leu Glu Gly Ile Phe Cys His Phe Pro Pro Ser Leu Leu
85 90 95

Pro Leu Ala Val Thr Leu Tyr Val Ser Trp Leu Leu Tyr Leu Phe Leu
100 105 110

Ile Leu Gly Val Thr Ala Ala Lys Phe Phe Cys Pro Asn Leu Ser Ala
115 120 125

Ile Ser Thr Thr Leu Lys Leu Ser His Asn Val Ala Gly Val Thr Phe
130 135 140

Leu Ala Phe Gly Asn Gly Ala Pro Asp Ile Phe Ser Ala Leu Val Ala
145 150 155 160

Phe Ser Asp Pro His Thr Ala Gly Leu Ala Leu Gly Ala Leu Phe Gly
165 170 175

Ala Gly Val Leu Val Thr Thr Val Val Ala Gly Gly Ile Thr Ile Leu
180 185 190

His Pro Phe Met Ala Ala Ser Arg Pro Phe Phe Arg Asp Ile Val Phe
195 200 205

Tyr Met Val Ala Val Phe Leu Thr Phe Leu Met Leu Phe Arg Gly Arg
210 215 220

Val Thr Leu Ala Trp Ala Leu Gly Tyr Leu Gly Leu Tyr Val Phe Tyr
225 230 235 240

Val Val Thr Val Ile Leu Cys Thr Trp Ile Tyr Gln Arg Gln Arg Arg
245 250 255

Gly Ser Leu Phe Cys Pro Met Pro Val Thr Pro Glu Ile Leu Ser Asp
260 265 270

Ser Glu Glu Asp Arg Val Ser Ser Asn Thr Asn Ser Tyr Asp Tyr Gly
275 280 285

Asp Glu Tyr Arg Pro Leu Phe Phe Tyr Gln Glu Thr Thr Ala Gln Ile
290 295 300

Leu Val Arg Ala Leu Asn Pro Leu Asp Tyr Met Lys Trp Arg Arg Lys
305 310 315 320

Ser Ala Tyr Trp Lys Ala Leu Lys Val Phe Lys Leu Pro Val Glu Phe
325 330 335

Leu Leu Leu Leu Thr Val Pro Val Val Asp Pro Asp Lys Asp Asp Gln
340 345 350

Asn Trp Lys Arg Pro Leu Asn Cys Leu His Leu Val Ile Ser Pro Leu
355 360 365

Val Val Val Leu Thr Leu Gln Ser Gly Thr Tyr Gly Val Tyr Glu Ile

-continued

```
    370                 375                 380

Gly Gly Leu Val Pro Val Trp Val Val Val Val Ile Ala Gly Thr Ala
385                 390                 395                 400

Leu Ala Ser Val Thr Phe Phe Ala Thr Ser Asp Ser Gln Pro Pro Arg
                405                 410                 415

Leu His Trp Leu Phe Ala Phe Leu Gly Phe Leu Thr Ser Ala Leu Trp
            420                 425                 430

Ile Asn Ala Ala Ala Thr Glu Val Val Asn Ile Leu Arg Ser Leu Gly
        435                 440                 445

Val Val Phe Arg Leu Ser Asn Thr Val Leu Gly Leu Thr Leu Leu Ala
    450                 455                 460

Trp Gly Asn Ser Ile Gly Asp Ala Phe Ser Asp Phe Thr Leu Ala Arg
465                 470                 475                 480

Gln Gly Tyr Pro Arg Met Ala Phe Ser Ala Cys Phe Gly Gly Ile Ile
                485                 490                 495

Phe Asn Ile Leu Val Gly Val Gly Leu Gly Cys Leu Leu Gln Ile Ser
            500                 505                 510

Arg Ser His Thr Glu Val Lys Leu Glu Pro Asp Gly Leu Leu Val Trp
        515                 520                 525

Val Leu Ala Gly Ala Leu Gly Leu Ser Leu Val Phe Ser Leu Val Ser
    530                 535                 540

Val Pro Leu Gln Cys Phe Gln Leu Ser Arg Val Tyr Gly Phe Cys Leu
545                 550                 555                 560

Leu Leu Phe Tyr Leu Asn Phe Leu Val Val Ala Leu Leu Thr Glu Phe
                565                 570                 575

Gly Val Ile His Leu Lys Ser Met
            580


<210> SEQ ID NO 15
<211> LENGTH: 1755
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(1755)

<400> SEQUENCE: 15

atg gcc ggc aga agg ctg aat ctg cgc tgg gca ctg agt gtg ctt tgt       48
Met Ala Gly Arg Arg Leu Asn Leu Arg Trp Ala Leu Ser Val Leu Cys
 1               5                  10                  15

gtg ctg cta atg gcg gag aca gtg tct ggg act agg ggc tcg tct aca       96
Val Leu Leu Met Ala Glu Thr Val Ser Gly Thr Arg Gly Ser Ser Thr
            20                  25                  30

gga gct cac att agc ccc cag ttt cca gct tca ggt gtg aac cag acc      144
Gly Ala His Ile Ser Pro Gln Phe Pro Ala Ser Gly Val Asn Gln Thr
        35                  40                  45

ccc gtg gta gac tgc cgc aag gtg tgt ggc ctg aat gtc tct gac cgc      192
Pro Val Val Asp Cys Arg Lys Val Cys Gly Leu Asn Val Ser Asp Arg
    50                  55                  60

tgt gac ttc atc cgg acc aac cct gac tgc cac agt gat ggg ggg tac      240
Cys Asp Phe Ile Arg Thr Asn Pro Asp Cys His Ser Asp Gly Gly Tyr
 65                  70                  75                  80

ctg gac tac ctg gaa ggc atc ttc tgc cac ttc cct ccc agc ctc ctc      288
Leu Asp Tyr Leu Glu Gly Ile Phe Cys His Phe Pro Pro Ser Leu Leu
                85                  90                  95
```

-continued

```
cct ctg gct gtc act ctc tac gtt tcc tgg ctg ctc tac ctg ttt ctg      336
Pro Leu Ala Val Thr Leu Tyr Val Ser Trp Leu Leu Tyr Leu Phe Leu
            100                 105                 110

att ctg gga gtc acc gca gcc aag ttt ttc tgc ccc aac ttg tcg gcc      384
Ile Leu Gly Val Thr Ala Ala Lys Phe Phe Cys Pro Asn Leu Ser Ala
        115                 120                 125

att tct acc aca ctg aag ctc tcc cac aac gtg gca ggc gtc acc ttc      432
Ile Ser Thr Thr Leu Lys Leu Ser His Asn Val Ala Gly Val Thr Phe
    130                 135                 140

ctg gca ttt ggg aat ggt gca cct gac atc ttc agt gcc ctg gtg gcc      480
Leu Ala Phe Gly Asn Gly Ala Pro Asp Ile Phe Ser Ala Leu Val Ala
145                 150                 155                 160

ttc tct gac ccg cac aca gcc ggc ctg gcc ctt ggg gca ctg ttt ggc      528
Phe Ser Asp Pro His Thr Ala Gly Leu Ala Leu Gly Ala Leu Phe Gly
                165                 170                 175

gct ggc gtg ctg gtt acc aca gtg gtg gcc gga ggc att acc atc cta      576
Ala Gly Val Leu Val Thr Thr Val Val Ala Gly Gly Ile Thr Ile Leu
            180                 185                 190

cac ccc ttc atg gct gcc tcc agg ccc ttc ttc agg gac atc gtt ttc      624
His Pro Phe Met Ala Ala Ser Arg Pro Phe Phe Arg Asp Ile Val Phe
        195                 200                 205

tac atg gtg gct gtg ttc ctg acc ttc ctc atg ctc ttc cgt ggc agg      672
Tyr Met Val Ala Val Phe Leu Thr Phe Leu Met Leu Phe Arg Gly Arg
    210                 215                 220

gtc acc ctg gca tgg gct ctg ggt tac ctg ggc ttg tat gtg ttc tat      720
Val Thr Leu Ala Trp Ala Leu Gly Tyr Leu Gly Leu Tyr Val Phe Tyr
225                 230                 235                 240

gtg gtc act gtg att ctc tgc acc tgg atc tac caa cgg caa cgg aga      768
Val Val Thr Val Ile Leu Cys Thr Trp Ile Tyr Gln Arg Gln Arg Arg
                245                 250                 255

gga tct ctg ttc tgc ccc atg cca gtt act cca gag atc ctc tca gac      816
Gly Ser Leu Phe Cys Pro Met Pro Val Thr Pro Glu Ile Leu Ser Asp
            260                 265                 270

tcc gag gag gac cgg gta tct tct aat acc aac agc tat gac tac ggt      864
Ser Glu Glu Asp Arg Val Ser Ser Asn Thr Asn Ser Tyr Asp Tyr Gly
        275                 280                 285

gat gag tac cgg ccg ctg ttc ttc tac cag gag acc acg gct cag atc      912
Asp Glu Tyr Arg Pro Leu Phe Phe Tyr Gln Glu Thr Thr Ala Gln Ile
    290                 295                 300

ctg gtc cgg gcc ctc aat ccc ctg gat tac atg aag tgg aga agg aaa      960
Leu Val Arg Ala Leu Asn Pro Leu Asp Tyr Met Lys Trp Arg Arg Lys
305                 310                 315                 320

tca gca tac tgg aaa gcc ctc aag gtg ttc aag ctg cct gtg gag ttc     1008
Ser Ala Tyr Trp Lys Ala Leu Lys Val Phe Lys Leu Pro Val Glu Phe
                325                 330                 335

ctg ctg ctc ctc aca gtc ccc gtc gtg gac ccg gac aag gat gac cag     1056
Leu Leu Leu Leu Thr Val Pro Val Val Asp Pro Asp Lys Asp Asp Gln
            340                 345                 350

aac tgg aaa cgg ccc ctc aac tgt ctg cat ctg gtt atc agc ccc ctg     1104
Asn Trp Lys Arg Pro Leu Asn Cys Leu His Leu Val Ile Ser Pro Leu
        355                 360                 365

gtt gtg gtc ctg acc ctg cag tcg ggg acc tat ggt gtc tat gag ata     1152
Val Val Val Leu Thr Leu Gln Ser Gly Thr Tyr Gly Val Tyr Glu Ile
    370                 375                 380

ggc ggc ctc gtt ccc gtc tgg gtc gtg gtg gtg atc gca ggc aca gcc     1200
Gly Gly Leu Val Pro Val Trp Val Val Val Val Ile Ala Gly Thr Ala
385                 390                 395                 400
```

-continued

```
ttg gct tca gtg acc ttt ttt gcc aca tct gac agc cag ccc ccc agg     1248
Leu Ala Ser Val Thr Phe Phe Ala Thr Ser Asp Ser Gln Pro Pro Arg
                405                 410                 415

ctt cac tgg ctc ttt gct ttc ctg ggc ttt ctg acc agc gcc ctg tgg     1296
Leu His Trp Leu Phe Ala Phe Leu Gly Phe Leu Thr Ser Ala Leu Trp
            420                 425                 430

atc aac gcg gcc gcc aca gag gtg gtg aac atc ttg cgg tcc ctg ggt     1344
Ile Asn Ala Ala Ala Thr Glu Val Val Asn Ile Leu Arg Ser Leu Gly
        435                 440                 445

gtg gtc ttc cgg ctg agc aac act gtg ctg ggg ctc acg ctg ctg gcc     1392
Val Val Phe Arg Leu Ser Asn Thr Val Leu Gly Leu Thr Leu Leu Ala
    450                 455                 460

tgg ggg aac agc att gga gat gcc ttc tcg gat ttc aca ctg gct cgc     1440
Trp Gly Asn Ser Ile Gly Asp Ala Phe Ser Asp Phe Thr Leu Ala Arg
465                 470                 475                 480

cag ggc tac cca cgg atg gcg ttc tcc gcc tgc ttt ggc ggc atc atc     1488
Gln Gly Tyr Pro Arg Met Ala Phe Ser Ala Cys Phe Gly Gly Ile Ile
                485                 490                 495

ttc aac atc ctc gtg ggt gtg ggg ctg ggc tgc ctg ctc cag atc tcc     1536
Phe Asn Ile Leu Val Gly Val Gly Leu Gly Cys Leu Leu Gln Ile Ser
            500                 505                 510

cga agc cac aca gaa gtg aag ctg gag cca gac gga ctg ctg gtg tgg     1584
Arg Ser His Thr Glu Val Lys Leu Glu Pro Asp Gly Leu Leu Val Trp
        515                 520                 525

gtc ctg gca ggc gcc ctg ggg ctc agc ctc gtc ttc tcc ctg gtc tca     1632
Val Leu Ala Gly Ala Leu Gly Leu Ser Leu Val Phe Ser Leu Val Ser
    530                 535                 540

gtc cca ttg cag tgc ttc cag ctc agc aga gtc tat ggc ttc tgc ctg     1680
Val Pro Leu Gln Cys Phe Gln Leu Ser Arg Val Tyr Gly Phe Cys Leu
545                 550                 555                 560

ctc ctc ttc tac ctg aac ttc ctt gtc gtg gcc ctc ctc act gaa ttt     1728
Leu Leu Phe Tyr Leu Asn Phe Leu Val Val Ala Leu Leu Thr Glu Phe
                565                 570                 575

gga gtg att cac ctg aaa agc atg tga                                 1755
Gly Val Ile His Leu Lys Ser Met  *
            580


<210> SEQ ID NO 16
<211> LENGTH: 152
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 16

Ile Leu Ile Val Leu Gly Ala Asp Leu Phe Val Asp Gly Ala Ser Ala
 1               5                  10                  15

Ile Ala Glu Val Leu Gly Ile Ser Glu Ser Val Ile Gly Leu Thr Leu
            20                  25                  30

Val Ala Leu Gly Thr Ser Leu Pro Glu Leu Phe Ala Ser Leu Ile Ala
        35                  40                  45

Ala Leu Lys Gly Gln Phe Gln Ala Asp Ile Ala Ile Gly Asn Val Ile
    50                  55                  60

Gly Ser Asn Ile Phe Asn Ile Leu Leu Gly Leu Gly Ile Ala Ser Leu
65                  70                  75                  80

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

-continued

```
Pro Ile Ser Leu Arg Arg Asp Val Leu Phe Leu Leu Leu Val Leu Leu
            100                 105                 110

Ile Leu Ile Val Phe Leu Leu Leu Gly Arg Ser Leu Ile Gly Arg Gly
        115                 120                 125

Asp Gly Val Leu Leu Leu Ile Leu Tyr Ile Leu Tyr Leu Thr Phe Leu
    130                 135                 140

Val Phe Ser Ile Leu Leu Glu Val
145                 150


<210> SEQ ID NO 17
<211> LENGTH: 661
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 17

Met Asp Leu Gln Gln Ser Thr Thr Ile Thr Ser Leu Glu Lys Trp Cys
 1               5                  10                  15

Leu Asp Glu Ser Leu Ser Gly Cys Arg Arg His Tyr Ser Val Lys Lys
            20                  25                  30

Lys Leu Lys Leu Ile Arg Val Leu Gly Leu Phe Met Gly Leu Val Ala
        35                  40                  45

Ile Ser Thr Val Ser Phe Ser Ile Ser Ala Phe Ser Glu Thr Asp Thr
    50                  55                  60

Gln Ser Thr Gly Glu Ala Ser Val Val Ser Gly Pro Arg Val Ala Gln
65                  70                  75                  80

Gly Tyr His Gln Arg Thr Leu Leu Asp Leu Asn Asp Lys Ile Leu Asp
                85                  90                  95

Tyr Thr Pro Gln Pro Pro Leu Ser Lys Glu Gly Glu Ser Glu Asn Ser
            100                 105                 110

Thr Asp His Ala Gln Gly Asp Tyr Pro Lys Asp Ile Phe Ser Leu Glu
        115                 120                 125

Glu Arg Arg Lys Gly Ala Ile Ile Leu His Val Ile Gly Met Ile Tyr
    130                 135                 140

Met Phe Ile Ala Leu Ala Ile Val Cys Asp Glu Phe Phe Val Pro Ser
145                 150                 155                 160

Leu Thr Val Ile Thr Glu Lys Leu Gly Ile Ser Asp Asp Val Ala Gly
                165                 170                 175

Ala Thr Phe Met Ala Ala Gly Gly Ser Ala Pro Glu Leu Phe Thr Ser
            180                 185                 190

Leu Ile Gly Val Phe Ile Ala His Ser Asn Val Gly Ile Gly Thr Ile
        195                 200                 205

Val Gly Ser Ala Val Phe Asn Ile Leu Phe Val Ile Gly Met Cys Ala
    210                 215                 220

Leu Phe Ser Arg Glu Ile Leu Asn Leu Thr Trp Trp Pro Leu Phe Arg
225                 230                 235                 240

Asp Val Ser Phe Tyr Ile Val Asp Leu Ile Met Leu Ile Ile Phe Phe
                245                 250                 255

Leu Asp Asn Val Ile Met Trp Trp Glu Ser Leu Leu Leu Leu Thr Ala
            260                 265                 270

Tyr Phe Cys Tyr Val Val Phe Met Lys Phe Asn Val Gln Val Glu Lys
        275                 280                 285
```

-continued

```
Trp Val Lys Gln Met Ile Asn Arg Asn Lys Val Val Lys Val Thr Ala
    290                 295                 300

Pro Glu Ala Gln Ala Lys Pro Ser Ala Ala Arg Asp Lys Asp Glu Pro
305                 310                 315                 320

Thr Leu Pro Ala Lys Pro Arg Leu Gln Arg Gly Gly Ser Ser Ala Ser
                325                 330                 335

Leu His Asn Ser Leu Met Arg Asn Ser Ile Phe Gln Leu Met Ile His
            340                 345                 350

Thr Leu Asp Pro Leu Ala Glu Glu Leu Gly Ser Tyr Gly Lys Leu Lys
        355                 360                 365

Tyr Tyr Asp Thr Met Thr Glu Glu Gly Arg Phe Arg Glu Lys Ala Ser
    370                 375                 380

Ile Leu His Lys Ile Ala Lys Lys Lys Cys His Val Asp Glu Asn Glu
385                 390                 395                 400

Arg Gln Asn Gly Ala Ala Asn His Val Glu Lys Ile Glu Leu Pro Asn
                405                 410                 415

Ser Thr Ser Thr Asp Val Glu Met Thr Pro Ser Ser Asp Ala Ser Glu
            420                 425                 430

Pro Val Gln Asn Gly Asn Leu Ser His Asn Ile Glu Gly Ala Glu Ala
        435                 440                 445

Gln Thr Ala Asp Glu Glu Glu Asp Gln Pro Leu Ser Leu Ala Trp Pro
    450                 455                 460

Ser Glu Thr Arg Lys Gln Val Thr Phe Leu Ile Val Phe Pro Ile Val
465                 470                 475                 480

Phe Pro Leu Trp Ile Thr Leu Pro Asp Val Arg Lys Pro Ser Ser Arg
                485                 490                 495

Lys Phe Phe Pro Ile Thr Phe Phe Gly Ser Ile Thr Trp Ile Ala Val
            500                 505                 510

Phe Ser Tyr Leu Met Val Trp Trp Ala His Gln Val Gly Glu Thr Ile
        515                 520                 525

Gly Ile Ser Glu Glu Ile Met Gly Leu Thr Ile Leu Ala Ala Gly Thr
    530                 535                 540

Ser Ile Pro Asp Leu Ile Thr Ser Val Ile Val Ala Arg Lys Gly Leu
545                 550                 555                 560

Gly Asp Met Ala Val Ser Ser Ser Val Gly Ser Asn Ile Phe Asp Ile
                565                 570                 575

Thr Val Gly Leu Pro Leu Pro Trp Leu Leu Tyr Thr Val Ile His Arg
            580                 585                 590

Phe Gln Pro Val Ala Val Ser Ser Asn Gly Leu Phe Cys Ala Ile Val
        595                 600                 605

Leu Leu Phe Ile Met Leu Leu Phe Val Ile Leu Ser Ile Ala Leu Cys
    610                 615                 620

Lys Trp Arg Met Asn Lys Ile Leu Gly Phe Ile Met Phe Gly Leu Tyr
625                 630                 635                 640

Phe Val Phe Leu Val Val Ser Val Leu Leu Glu Asp Arg Ile Leu Thr
                645                 650                 655

Cys Pro Val Ser Ile
            660
```

<210> SEQ ID NO 18
<211> LENGTH: 1408
<212> TYPE: DNA

-continued

```
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (84)...(1343)

<400> SEQUENCE: 18

ccacgcgtcc gcggacgcgt gggtcggcac cggcgaggcc gtgctggaac ccgggcctca     60

gccgcagccg cagcggggcc gac atg acg aca gct ccc cag gag ccc ccc gcc    113
Met Thr Thr Ala Pro Gln Glu Pro Pro Ala
 1               5                  10

cgg ccc ctc cag gcg ggc agt gga gct ggc ccg gcg cct ggg cgc gcc      161
Arg Pro Leu Gln Ala Gly Ser Gly Ala Gly Pro Ala Pro Gly Arg Ala
                15                  20                  25

atg cgc agc acc acg ctc ctg gcc ctg ctg gcg ctg gtc ttg ctt tac      209
Met Arg Ser Thr Thr Leu Leu Ala Leu Leu Ala Leu Val Leu Leu Tyr
            30                  35                  40

ttg gtg tct ggt gcc ctg gtg ttc cgg gcc ctg gag cag ccc cac gag      257
Leu Val Ser Gly Ala Leu Val Phe Arg Ala Leu Glu Gln Pro His Glu
        45                  50                  55

cag cag gcc cag agg gag ctg ggg gag gtc cga gag aag ttc ctg agg      305
Gln Gln Ala Gln Arg Glu Leu Gly Glu Val Arg Glu Lys Phe Leu Arg
    60                  65                  70

gcc cat ccg tgt gtg agc gac cag gag ctg ggc ctc ctc atc aag gag      353
Ala His Pro Cys Val Ser Asp Gln Glu Leu Gly Leu Leu Ile Lys Glu
 75                  80                  85                  90

gtg gct gat gcc ctg gga ggg ggt gcg gac cca gaa acc aac tcg acc      401
Val Ala Asp Ala Leu Gly Gly Gly Ala Asp Pro Glu Thr Asn Ser Thr
                95                 100                 105

agc aac agc agc cac tca gcc tgg gac ctg ggc agc gcc ttc ttt ttc      449
Ser Asn Ser Ser His Ser Ala Trp Asp Leu Gly Ser Ala Phe Phe Phe
            110                 115                 120

tca ggg acc atc atc acc acc atc ggc tat ggc aat gtg gcc ctg cgc      497
Ser Gly Thr Ile Ile Thr Thr Ile Gly Tyr Gly Asn Val Ala Leu Arg
        125                 130                 135

aca gat gcc ggg cgc ctc ttc tgc atc ttt tat gcg ctg gtg ggg att      545
Thr Asp Ala Gly Arg Leu Phe Cys Ile Phe Tyr Ala Leu Val Gly Ile
    140                 145                 150

ccg ctg ttt ggg atc cta ctg gca ggg gtc ggg gac cgg ctg ggc tcc      593
Pro Leu Phe Gly Ile Leu Leu Ala Gly Val Gly Asp Arg Leu Gly Ser
155                 160                 165                 170

tcc ctg cgc cat ggc atc ggt cac att gaa gcc atc ttc ttg aag tgg      641
Ser Leu Arg His Gly Ile Gly His Ile Glu Ala Ile Phe Leu Lys Trp
                175                 180                 185

cac gtg cca ccg gag cta gta aga gtg ctg tcg gcg atg ctt ttc ctg      689
His Val Pro Pro Glu Leu Val Arg Val Leu Ser Ala Met Leu Phe Leu
            190                 195                 200

ctg atc ggc tgc ctg ctc ttt gtc ctc acg ccc acg ttc gtg ttc tgc      737
Leu Ile Gly Cys Leu Leu Phe Val Leu Thr Pro Thr Phe Val Phe Cys
        205                 210                 215

tat atg gag gac tgg agc aag ctg gag gcc atc tac ttt gtc ata gtg      785
Tyr Met Glu Asp Trp Ser Lys Leu Glu Ala Ile Tyr Phe Val Ile Val
    220                 225                 230

acg ctt acc acc gtg ggc ttt ggc gac tat gtg gcc ggc gcg gac ccc      833
Thr Leu Thr Thr Val Gly Phe Gly Asp Tyr Val Ala Gly Ala Asp Pro
235                 240                 245                 250

agg cag gac tcc ccg gcc tat cag ccg ctg gtg tgg ttc tgg atc ctg      881
Arg Gln Asp Ser Pro Ala Tyr Gln Pro Leu Val Trp Phe Trp Ile Leu
                255                 260                 265
```

-continued

```
ctc ggc ctg gct tac ttc gcc tca gtg ctc acc acc atc ggg aac tgg      929
Leu Gly Leu Ala Tyr Phe Ala Ser Val Leu Thr Thr Ile Gly Asn Trp
            270                 275                 280

ctg cga gta gtg tcc cgc cgc act cgg gca gag atg ggc ggc ctc acg      977
Leu Arg Val Val Ser Arg Arg Thr Arg Ala Glu Met Gly Gly Leu Thr
        285                 290                 295

gct cag gct gcc agc tgg act ggc aca gtg aca gcg cgc gtg acc cag     1025
Ala Gln Ala Ala Ser Trp Thr Gly Thr Val Thr Ala Arg Val Thr Gln
    300                 305                 310

cga gcc ggg ccc gcc gcc ccg ccg ccg gag aag gag cag cca ctg ctg     1073
Arg Ala Gly Pro Ala Ala Pro Pro Pro Glu Lys Glu Gln Pro Leu Leu
315                 320                 325                 330

cct cca ccg ccc tgt cca gcg cag ccg ctg ggc agg ccc cga tcc cct     1121
Pro Pro Pro Pro Cys Pro Ala Gln Pro Leu Gly Arg Pro Arg Ser Pro
                335                 340                 345

tcg ccc ccc gag aag gct cag ccg cct tcc ccg ccc acg gcc tcg gcc     1169
Ser Pro Pro Glu Lys Ala Gln Pro Pro Ser Pro Pro Thr Ala Ser Ala
            350                 355                 360

ctg gat tat ccc agc gag aac ctg gcc ttc atc gac gag tcc tcg gat     1217
Leu Asp Tyr Pro Ser Glu Asn Leu Ala Phe Ile Asp Glu Ser Ser Asp
        365                 370                 375

acg cag agc gag cgc ggc tgc ccg ctg ccc cgc gcg ccg aga ggt cgc     1265
Thr Gln Ser Glu Arg Gly Cys Pro Leu Pro Arg Ala Pro Arg Gly Arg
    380                 385                 390

cgc cgc cca aat ccc ccc agg aag ccc gtg cgg ccc cgc ggc ccc ggg     1313
Arg Arg Pro Asn Pro Pro Arg Lys Pro Val Arg Pro Arg Gly Pro Gly
395                 400                 405                 410

cgt ccc cga gac aaa ggc gtg ccg gtg tag gggcaggatc cctggccggg       1363
Arg Pro Arg Asp Lys Gly Val Pro Val  *
                415

cctctcaagg gcttcgtttc tgctctcccc ggcatgcctg gcttc                   1408


<210> SEQ ID NO 19
<211> LENGTH: 419
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 19

Met Thr Thr Ala Pro Gln Glu Pro Pro Ala Arg Pro Leu Gln Ala Gly
 1               5                  10                  15

Ser Gly Ala Gly Pro Ala Pro Gly Arg Ala Met Arg Ser Thr Thr Leu
            20                  25                  30

Leu Ala Leu Leu Ala Leu Val Leu Leu Tyr Leu Val Ser Gly Ala Leu
        35                  40                  45

Val Phe Arg Ala Leu Glu Gln Pro His Glu Gln Gln Ala Gln Arg Glu
    50                  55                  60

Leu Gly Glu Val Arg Glu Lys Phe Leu Arg Ala His Pro Cys Val Ser
65                  70                  75                  80

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

Gly Gly Ala Asp Pro Glu Thr Asn Ser Thr Ser Asn Ser Ser His Ser
            100                 105                 110

Ala Trp Asp Leu Gly Ser Ala Phe Phe Phe Ser Gly Thr Ile Ile Thr
        115                 120                 125

Thr Ile Gly Tyr Gly Asn Val Ala Leu Arg Thr Asp Ala Gly Arg Leu
```

-continued

```
    130                 135                 140

Phe Cys Ile Phe Tyr Ala Leu Val Gly Ile Pro Leu Phe Gly Ile Leu
145                 150                 155                 160

Leu Ala Gly Val Gly Asp Arg Leu Gly Ser Ser Leu Arg His Gly Ile
                165                 170                 175

Gly His Ile Glu Ala Ile Phe Leu Lys Trp His Val Pro Pro Glu Leu
            180                 185                 190

Val Arg Val Leu Ser Ala Met Leu Phe Leu Leu Ile Gly Cys Leu Leu
        195                 200                 205

Phe Val Leu Thr Pro Thr Phe Val Phe Cys Tyr Met Glu Asp Trp Ser
    210                 215                 220

Lys Leu Glu Ala Ile Tyr Phe Val Ile Val Thr Leu Thr Thr Val Gly
225                 230                 235                 240

Phe Gly Asp Tyr Val Ala Gly Ala Asp Pro Arg Gln Asp Ser Pro Ala
                245                 250                 255

Tyr Gln Pro Leu Val Trp Phe Trp Ile Leu Leu Gly Leu Ala Tyr Phe
            260                 265                 270

Ala Ser Val Leu Thr Thr Ile Gly Asn Trp Leu Arg Val Val Ser Arg
        275                 280                 285

Arg Thr Arg Ala Glu Met Gly Gly Leu Thr Ala Gln Ala Ala Ser Trp
    290                 295                 300

Thr Gly Thr Val Thr Ala Arg Val Thr Gln Arg Ala Gly Pro Ala Ala
305                 310                 315                 320

Pro Pro Pro Glu Lys Glu Gln Pro Leu Leu Pro Pro Pro Pro Cys Pro
                325                 330                 335

Ala Gln Pro Leu Gly Arg Pro Arg Ser Pro Ser Pro Pro Glu Lys Ala
            340                 345                 350

Gln Pro Pro Ser Pro Pro Thr Ala Ser Ala Leu Asp Tyr Pro Ser Glu
        355                 360                 365

Asn Leu Ala Phe Ile Asp Glu Ser Ser Asp Thr Gln Ser Glu Arg Gly
    370                 375                 380

Cys Pro Leu Pro Arg Ala Pro Arg Gly Arg Arg Arg Pro Asn Pro Pro
385                 390                 395                 400

Arg Lys Pro Val Arg Pro Arg Gly Pro Gly Arg Pro Arg Asp Lys Gly
                405                 410                 415

Val Pro Val


<210> SEQ ID NO 20
<211> LENGTH: 1260
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(1260)

<400> SEQUENCE: 20

atg acg aca gct ccc cag gag ccc ccc gcc cgg ccc ctc cag gcg ggc       48
Met Thr Thr Ala Pro Gln Glu Pro Pro Ala Arg Pro Leu Gln Ala Gly
 1               5                  10                  15

agt gga gct ggc ccg gcg cct ggg cgc gcc atg cgc agc acc acg ctc       96
Ser Gly Ala Gly Pro Ala Pro Gly Arg Ala Met Arg Ser Thr Thr Leu
            20                  25                  30

ctg gcc ctg ctg gcg ctg gtc ttg ctt tac ttg gtg tct ggt gcc ctg      144
Leu Ala Leu Leu Ala Leu Val Leu Leu Tyr Leu Val Ser Gly Ala Leu
```

-continued

```
         35                  40                  45

gtg ttc cgg gcc ctg gag cag ccc cac gag cag cag gcc cag agg gag      192
Val Phe Arg Ala Leu Glu Gln Pro His Glu Gln Gln Ala Gln Arg Glu
     50                  55                  60

ctg ggg gag gtc cga gag aag ttc ctg agg gcc cat ccg tgt gtg agc      240
Leu Gly Glu Val Arg Glu Lys Phe Leu Arg Ala His Pro Cys Val Ser
 65                  70                  75                  80

gac cag gag ctg ggc ctc ctc atc aag gag gtg gct gat gcc ctg gga      288
Asp Gln Glu Leu Gly Leu Leu Ile Lys Glu Val Ala Asp Ala Leu Gly
                 85                  90                  95

ggg ggt gcg gac cca gaa acc aac tcg acc agc aac agc agc cac tca      336
Gly Gly Ala Asp Pro Glu Thr Asn Ser Thr Ser Asn Ser Ser His Ser
            100                 105                 110

gcc tgg gac ctg ggc agc gcc ttc ttt ttc tca ggg acc atc atc acc      384
Ala Trp Asp Leu Gly Ser Ala Phe Phe Phe Ser Gly Thr Ile Ile Thr
        115                 120                 125

acc atc ggc tat ggc aat gtg gcc ctg cgc aca gat gcc ggg cgc ctc      432
Thr Ile Gly Tyr Gly Asn Val Ala Leu Arg Thr Asp Ala Gly Arg Leu
    130                 135                 140

ttc tgc atc ttt tat gcg ctg gtg ggg att ccg ctg ttt ggg atc cta      480
Phe Cys Ile Phe Tyr Ala Leu Val Gly Ile Pro Leu Phe Gly Ile Leu
145                 150                 155                 160

ctg gca ggg gtc ggg gac cgg ctg ggc tcc tcc ctg cgc cat ggc atc      528
Leu Ala Gly Val Gly Asp Arg Leu Gly Ser Ser Leu Arg His Gly Ile
                165                 170                 175

ggt cac att gaa gcc atc ttc ttg aag tgg cac gtg cca ccg gag cta      576
Gly His Ile Glu Ala Ile Phe Leu Lys Trp His Val Pro Pro Glu Leu
            180                 185                 190

gta aga gtg ctg tcg gcg atg ctt ttc ctg ctg atc ggc tgc ctg ctc      624
Val Arg Val Leu Ser Ala Met Leu Phe Leu Leu Ile Gly Cys Leu Leu
        195                 200                 205

ttt gtc ctc acg ccc acg ttc gtg ttc tgc tat atg gag gac tgg agc      672
Phe Val Leu Thr Pro Thr Phe Val Phe Cys Tyr Met Glu Asp Trp Ser
    210                 215                 220

aag ctg gag gcc atc tac ttt gtc ata gtg acg ctt acc acc gtg ggc      720
Lys Leu Glu Ala Ile Tyr Phe Val Ile Val Thr Leu Thr Thr Val Gly
225                 230                 235                 240

ttt ggc gac tat gtg gcc ggc gcg gac ccc agg cag gac tcc ccg gcc      768
Phe Gly Asp Tyr Val Ala Gly Ala Asp Pro Arg Gln Asp Ser Pro Ala
                245                 250                 255

tat cag ccg ctg gtg tgg ttc tgg atc ctg ctc ggc ctg gct tac ttc      816
Tyr Gln Pro Leu Val Trp Phe Trp Ile Leu Leu Gly Leu Ala Tyr Phe
            260                 265                 270

gcc tca gtg ctc acc acc atc ggg aac tgg ctg cga gta gtg tcc cgc      864
Ala Ser Val Leu Thr Thr Ile Gly Asn Trp Leu Arg Val Val Ser Arg
        275                 280                 285

cgc act cgg gca gag atg ggc ggc ctc acg gct cag gct gcc agc tgg      912
Arg Thr Arg Ala Glu Met Gly Gly Leu Thr Ala Gln Ala Ala Ser Trp
    290                 295                 300

act ggc aca gtg aca gcg cgc gtg acc cag cga gcc ggg ccc gcc gcc      960
Thr Gly Thr Val Thr Ala Arg Val Thr Gln Arg Ala Gly Pro Ala Ala
305                 310                 315                 320

ccg ccg ccg gag aag gag cag cca ctg ctg cct cca ccg ccc tgt cca     1008
Pro Pro Pro Glu Lys Glu Gln Pro Leu Leu Pro Pro Pro Pro Cys Pro
                325                 330                 335

gcg cag ccg ctg ggc agg ccc cga tcc cct tcg ccc ccc gag aag gct     1056
Ala Gln Pro Leu Gly Arg Pro Arg Ser Pro Ser Pro Pro Glu Lys Ala
```

-continued

```
            340                 345                 350

cag ccg cct tcc ccg ccc acg gcc tcg gcc ctg gat tat ccc agc gag     1104
Gln Pro Pro Ser Pro Pro Thr Ala Ser Ala Leu Asp Tyr Pro Ser Glu
        355                 360                 365

aac ctg gcc ttc atc gac gag tcc tcg gat acg cag agc gag cgc ggc     1152
Asn Leu Ala Phe Ile Asp Glu Ser Ser Asp Thr Gln Ser Glu Arg Gly
    370                 375                 380

tgc ccg ctg ccc cgc gcg ccg aga ggt cgc cgc cgc cca aat ccc ccc     1200
Cys Pro Leu Pro Arg Ala Pro Arg Gly Arg Arg Arg Pro Asn Pro Pro
385                 390                 395                 400

agg aag ccc gtg cgg ccc cgc ggc ccc ggg cgt ccc cga gac aaa ggc     1248
Arg Lys Pro Val Arg Pro Arg Gly Pro Gly Arg Pro Arg Asp Lys Gly
                405                 410                 415

gtg ccg gtg tag                                                     1260
Val Pro Val  *


<210> SEQ ID NO 21
<211> LENGTH: 4586
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (146)...(2518)

<400> SEQUENCE: 21

ccacgcgtcc gattccagcc atccctctgc ctgcaatgag agcttcccgc cgcctcagcc     60

acagtcccac ccgggggcct tgggccccag acatgcggtg atctcagggc aagggttgcc    120

acgaccaccc agaacctcac cagcc atg aaa gcc cac ccc aag gag atg gtg      172
Met Lys Ala His Pro Lys Glu Met Val
 1               5

cct ctc atg ggc aag aga gtt gct gcc ccc agt ggg aac cct gcc gtc      220
Pro Leu Met Gly Lys Arg Val Ala Ala Pro Ser Gly Asn Pro Ala Val
 10                  15                  20                  25

ctg cca gag aag agg ccg gcg gag atc acc ccc aca aag aag agt gca      268
Leu Pro Glu Lys Arg Pro Ala Glu Ile Thr Pro Thr Lys Lys Ser Ala
                30                  35                  40

cac ttc ttc ctg gag ata gaa ggg ttt gaa ccc aac ccc aca gtt gcc      316
His Phe Phe Leu Glu Ile Glu Gly Phe Glu Pro Asn Pro Thr Val Ala
            45                  50                  55

aag acc tct cct cct gtc ttc tcc aag ccc atg gat tcc aac atc cgg      364
Lys Thr Ser Pro Pro Val Phe Ser Lys Pro Met Asp Ser Asn Ile Arg
        60                  65                  70

cag tgc atc tct ggt aac tgt gat gac atg gac tcc ccc cag tct cct      412
Gln Cys Ile Ser Gly Asn Cys Asp Asp Met Asp Ser Pro Gln Ser Pro
    75                  80                  85

caa gat gat gtg aca gag acc cca tcc aat ccc aac agc ccc agt gca      460
Gln Asp Asp Val Thr Glu Thr Pro Ser Asn Pro Asn Ser Pro Ser Ala
 90                  95                 100                 105

cag ctg gcc aag gaa gag cag agg agg aaa aag agg cgg ctg aag aag      508
Gln Leu Ala Lys Glu Glu Gln Arg Arg Lys Lys Arg Arg Leu Lys Lys
                110                 115                 120

cgc atc ttt gca gcc gtg tct gag ggc tgc gtg gag gag ttg gta gag      556
Arg Ile Phe Ala Ala Val Ser Glu Gly Cys Val Glu Glu Leu Val Glu
            125                 130                 135

ttg ctg gtg gag ctg cag gag ctt tgc agg cgg cgc cat gat gag gat      604
Leu Leu Val Glu Leu Gln Glu Leu Cys Arg Arg Arg His Asp Glu Asp
        140                 145                 150
```

-continued

```
gtg cct gac ttc ctc atg cac aag ctg acg gcc tcc gac acg ggg aag      652
Val Pro Asp Phe Leu Met His Lys Leu Thr Ala Ser Asp Thr Gly Lys
    155                 160                 165

acc tgc ctg atg aag gcc ttg tta aac atc aac ccc aac acc aag gag      700
Thr Cys Leu Met Lys Ala Leu Leu Asn Ile Asn Pro Asn Thr Lys Glu
170                 175                 180                 185

atc gtg cgg atc ctg ctt gcc ttt gct gaa gag aac gac atc ctg ggc      748
Ile Val Arg Ile Leu Leu Ala Phe Ala Glu Glu Asn Asp Ile Leu Gly
                190                 195                 200

agg ttc atc aac gcc gag tac aca gag gag gcc tat gaa ggg cag acg      796
Arg Phe Ile Asn Ala Glu Tyr Thr Glu Glu Ala Tyr Glu Gly Gln Thr
            205                 210                 215

gcg ctg aac atc gcc atc gag cgg cgg cag ggg gac atc gca gcc ctg      844
Ala Leu Asn Ile Ala Ile Glu Arg Arg Gln Gly Asp Ile Ala Ala Leu
        220                 225                 230

ctc atc gcc gcc ggc gcc gac gtc aac gcg cac gcc aag ggg gcc ttc      892
Leu Ile Ala Ala Gly Ala Asp Val Asn Ala His Ala Lys Gly Ala Phe
    235                 240                 245

ttc aac ccc aag tac caa cac gaa ggc ttc tac ttc ggt gag acg ccc      940
Phe Asn Pro Lys Tyr Gln His Glu Gly Phe Tyr Phe Gly Glu Thr Pro
250                 255                 260                 265

ctg gcc ctg gca gca tgc acc aac cag ccc gag att gtg cag ctg ctg      988
Leu Ala Leu Ala Ala Cys Thr Asn Gln Pro Glu Ile Val Gln Leu Leu
                270                 275                 280

atg gag cac gag cag acg gac atc acc tcg cgg gac tca cga ggc aac     1036
Met Glu His Glu Gln Thr Asp Ile Thr Ser Arg Asp Ser Arg Gly Asn
            285                 290                 295

aac atc ctt cac gcc ctg gtg acc gtg gcc gag gac ttc aag acg cag     1084
Asn Ile Leu His Ala Leu Val Thr Val Ala Glu Asp Phe Lys Thr Gln
        300                 305                 310

aat gac ttt gtg aag cgc atg tac gac atg atc cta ctg cgg agt ggc     1132
Asn Asp Phe Val Lys Arg Met Tyr Asp Met Ile Leu Leu Arg Ser Gly
    315                 320                 325

aac tgg gag ctg gag acc act cgc aac aac gat ggc ctc acg ccg ctg     1180
Asn Trp Glu Leu Glu Thr Thr Arg Asn Asn Asp Gly Leu Thr Pro Leu
330                 335                 340                 345

cag ctg gcc gcc aag atg ggc aag gcg gag atc ctg aag tac atc ctc     1228
Gln Leu Ala Ala Lys Met Gly Lys Ala Glu Ile Leu Lys Tyr Ile Leu
                350                 355                 360

agt cgt gag atc aag gag aag cgg ctc cgg agc ctg tcc agg aag ttc     1276
Ser Arg Glu Ile Lys Glu Lys Arg Leu Arg Ser Leu Ser Arg Lys Phe
            365                 370                 375

acc gac tgg gcg tac gga ccc gtg tca tcc tcc ctc tac gac ctc acc     1324
Thr Asp Trp Ala Tyr Gly Pro Val Ser Ser Ser Leu Tyr Asp Leu Thr
        380                 385                 390

aac gtg gac acc acc acg gac aac tca gtg ctg gaa atc act gtc tac     1372
Asn Val Asp Thr Thr Thr Asp Asn Ser Val Leu Glu Ile Thr Val Tyr
    395                 400                 405

aac acc aac atc gac aac cgg cat gag atg ctg acc ctg gag ccg ctg     1420
Asn Thr Asn Ile Asp Asn Arg His Glu Met Leu Thr Leu Glu Pro Leu
410                 415                 420                 425

cac acg ctg ctg cat atg aag tgg aag aag ttt gcc aag cac atg ttc     1468
His Thr Leu Leu His Met Lys Trp Lys Lys Phe Ala Lys His Met Phe
                430                 435                 440

ttt ctg tcc ttc tgc ttt tat ttc ttc tac aac atc acc ctg acc ctc     1516
Phe Leu Ser Phe Cys Phe Tyr Phe Phe Tyr Asn Ile Thr Leu Thr Leu
            445                 450                 455
```

-continued

```
gtc tcg tac tac cgc ccc cgg gag gag gag gcc atc ccg cac ccc ttg     1564
Val Ser Tyr Tyr Arg Pro Arg Glu Glu Glu Ala Ile Pro His Pro Leu
        460                 465                 470

gcc ctg acg cac aag atg ggg tgg ctg cag ctc cta ggg agg atg ttt     1612
Ala Leu Thr His Lys Met Gly Trp Leu Gln Leu Leu Gly Arg Met Phe
    475                 480                 485

gtg ctc atc tgg gcc atg tgc atc tct gtg aaa gag ggc att gcc atc     1660
Val Leu Ile Trp Ala Met Cys Ile Ser Val Lys Glu Gly Ile Ala Ile
490                 495                 500                 505

ttc ctg ctg aga ccc tcg gat ctg cag tcc atc ctc tcg gat gcc tgg     1708
Phe Leu Leu Arg Pro Ser Asp Leu Gln Ser Ile Leu Ser Asp Ala Trp
                510                 515                 520

ttc cac ttt gtc ttt ttt atc caa gct gtg ctt gtg ata ctg tct gtc     1756
Phe His Phe Val Phe Phe Ile Gln Ala Val Leu Val Ile Leu Ser Val
            525                 530                 535

ttc ttg tac ttg ttt gcc tac aaa gag tac ctc gcc tgc ctc gtg ctg     1804
Phe Leu Tyr Leu Phe Ala Tyr Lys Glu Tyr Leu Ala Cys Leu Val Leu
        540                 545                 550

gcc atg gcc ctg ggc tgg gcg aac atg ctc tac tat acg cgg ggt ttc     1852
Ala Met Ala Leu Gly Trp Ala Asn Met Leu Tyr Tyr Thr Arg Gly Phe
    555                 560                 565

cag tcc atg ggc atg tac agc gtc atg atc cag aag gtc att ttg cat     1900
Gln Ser Met Gly Met Tyr Ser Val Met Ile Gln Lys Val Ile Leu His
570                 575                 580                 585

gat gtt ctg aag ttc ttg ttt gta tat atc gtg ttt ttg ctt gga ttt     1948
Asp Val Leu Lys Phe Leu Phe Val Tyr Ile Val Phe Leu Leu Gly Phe
                590                 595                 600

gga gta gcc ttg gcc tcg ctg atc gag aag tgt ccc aaa gac aac aag     1996
Gly Val Ala Leu Ala Ser Leu Ile Glu Lys Cys Pro Lys Asp Asn Lys
            605                 610                 615

gac tgc agc tcc tac ggc agc ttc agt gac gca gtg ctg gaa ctc ttc     2044
Asp Cys Ser Ser Tyr Gly Ser Phe Ser Asp Ala Val Leu Glu Leu Phe
        620                 625                 630

aag ctc acc ata ggc ctg ggt gat ctg aac atc cag cag aac tcc aag     2092
Lys Leu Thr Ile Gly Leu Gly Asp Leu Asn Ile Gln Gln Asn Ser Lys
    635                 640                 645

tat ccc att ctc ttt ctg ttc ctg ctc atc acc tat gtc atc ctc acc     2140
Tyr Pro Ile Leu Phe Leu Phe Leu Leu Ile Thr Tyr Val Ile Leu Thr
650                 655                 660                 665

ttt gtt ctc ctc ctc aac atg ctc att gct ctg atg ggc gag act gtg     2188
Phe Val Leu Leu Leu Asn Met Leu Ile Ala Leu Met Gly Glu Thr Val
                670                 675                 680

gag aac gtc tcc aag gag agc gaa cgc atc tgg cgc ctg cag aga gcc     2236
Glu Asn Val Ser Lys Glu Ser Glu Arg Ile Trp Arg Leu Gln Arg Ala
            685                 690                 695

agg acc atc ttg gag ttt gag aaa atg tta cca gaa tgg ctg agg agc     2284
Arg Thr Ile Leu Glu Phe Glu Lys Met Leu Pro Glu Trp Leu Arg Ser
        700                 705                 710

aga ttc cgg atg gga gag ctg tgc aaa gtg gcc gag gat gat ttc cga     2332
Arg Phe Arg Met Gly Glu Leu Cys Lys Val Ala Glu Asp Asp Phe Arg
    715                 720                 725

ctg tgt ttg cgg atc aat gag gtg aag tgg act gaa tgg aag acg cac     2380
Leu Cys Leu Arg Ile Asn Glu Val Lys Trp Thr Glu Trp Lys Thr His
730                 735                 740                 745

gtc tcc ttc ctt aac gaa gac ccg ggg cct gta aga cga aca gat ttc     2428
Val Ser Phe Leu Asn Glu Asp Pro Gly Pro Val Arg Arg Thr Asp Phe
                750                 755                 760
```

-continued

```
aac aaa atc caa gat tct tcc agg aac aac agc aaa acc act ctc aat     2476
Asn Lys Ile Gln Asp Ser Ser Arg Asn Asn Ser Lys Thr Thr Leu Asn
            765                 770                 775

gca ttt gaa gaa gtc gag gaa ttc ccg gaa acc tcg gtg tag             2518
Ala Phe Glu Glu Val Glu Glu Phe Pro Glu Thr Ser Val  *
        780                 785                 790

aagcggaacc cagagctggt gtgcgcgtgc gctgtctggc gctgcaggcg gagtcaccga   2578

ctctgtgcag agaggctttg agggatgatg gagtccggct ctgctggcct agaagcagag   2638

tgcaccctcg tgctcagtgc tcagtgggtg tctgaactga ggggcagttg tcaatttgtc   2698

tgagtgggaa acatcctgga ttttgttact tggcaaacag ctggtgtaaa cctacagcca   2758

gcagcagtct ggagcctggg agcctcctga agtcccgggt gaagcctctg gttttaccaa   2818

ttgcaggtcg gcttggctgg gagagatgga tggcgggaaa ggggcagcag tcttgaggag   2878

cagggagagg agtctttcct cctgccagct tcccccgtca gccccaaccc cagcccacac   2938

attgtaccat ctcttctgct gtgactgggt tgcctgaatt tgtgggagac ccgtgatccc   2998

atcccagagt gtgcgggggga cggaggtaag ctggatatcc tgggggagga ggggaatgcg   3058

ctctggaaac acccttccgg aacccttcgg ggaaaaggag accatccttg gagtgaacgt   3118

cccctgacac cccaaggttc aaactgtctc aagctgagag atgtttttag tagcataatt   3178

aacacagggt tttaacttgc aatacggaaa agacatttca gttgagaatg aaaattacta   3238

caatgaagtt tgtgatttta aaagtggaga cagactgggg gctttggggc tggatgtaag   3298

tattatatat ttggcctcag ggtgcccaga gcaagacaaa aagcttttct tcacacacac   3358

aaaagtctgc atgagacact ccgggcaagt cctgctgggc cgccgcgatc tgggtgaaag   3418

gtcctggctc ttttcctcgt cctgacctca cagtagcgca tgcctgtgtg ctgggatcgt   3478

ggctttcgct gaagcagaaa tagcagctgc tcgatcgata tcatcttgga actcagcagt   3538

tagtcgcata cctcagtacg tctcagtggg ggaatttaac aaaatgcctc aactgctttg   3598

gtacgaagta tttttttttt aattttaact gtgaattttg aagctgaagg ggaagcttgt   3658

gagagaaaag catttgccaa gactttgagc ttatttttag gtcctcgtcc tctgatgttc   3718

tctttctgaa atgacacgga gtcagtctcg ggggcagagg tgaagtggag acggaaggat   3778

tttccaggtg actggggccg aaaccaccag aaaatccact ctgccgccgt tatctggtga   3838

aaggattcat gtaaaaatgt tcgaggtgga attataaaaa tagtaaccat aaatgttaat   3898

cttaaatggc agaaatagaa atttggcctt cagataacat ggcgatagat aagttcatct   3958

ggcttgaggc aaactgaaga gtcggggcct agcattgcac tctgggccag tttctctgcc   4018

ctgggccact ctgtgtgcca gactagctgg acagatagag actttgtgcc cctgatgggg   4078

ccgattgggg agaggtgggc tggggtgtgc aggcttcaca atccacagca gcccctgccc   4138

tcccagctga cccagggagt aatcgcgtgc tctaagccac agtggtcggg gctgggcatg   4198

ggcctctgga gaagagaaga tttgaggaga actgtcctag aggcaggagg agcagatgtg   4258

tttcagaatg ggcagaatta ggaaattgag aaagattttg gctcaacaga atccagcaac   4318

tgctccagat gttggagatg tttaagcaga agctggttga gcacttaatg aggaatgttg   4378

ttgaaaatgg tcattggaag aagtttaagg tccctttag cctggagatt gtacaaatca   4438

gcattccaca tctggagtta gctacccgca ttaagcctga acagacatct tggtctgaaa   4498

ggaagtggtt tggattcatg atgccaagct ccacactatg gagctgggaa ttccagaatt   4558
```

-continued

```
gctttgactc agatattaat ggagaaag                                        4586


<210> SEQ ID NO 22
<211> LENGTH: 790
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 22

Met Lys Ala His Pro Lys Glu Met Val Pro Leu Met Gly Lys Arg Val
 1               5                  10                  15

Ala Ala Pro Ser Gly Asn Pro Ala Val Leu Pro Glu Lys Arg Pro Ala
            20                  25                  30

Glu Ile Thr Pro Thr Lys Lys Ser Ala His Phe Phe Leu Glu Ile Glu
        35                  40                  45

Gly Phe Glu Pro Asn Pro Thr Val Ala Lys Thr Ser Pro Pro Val Phe
    50                  55                  60

Ser Lys Pro Met Asp Ser Asn Ile Arg Gln Cys Ile Ser Gly Asn Cys
65                  70                  75                  80

Asp Asp Met Asp Ser Pro Gln Ser Pro Gln Asp Asp Val Thr Glu Thr
                85                  90                  95

Pro Ser Asn Pro Asn Ser Pro Ser Ala Gln Leu Ala Lys Glu Glu Gln
            100                 105                 110

Arg Arg Lys Lys Arg Arg Leu Lys Lys Arg Ile Phe Ala Ala Val Ser
        115                 120                 125

Glu Gly Cys Val Glu Glu Leu Val Glu Leu Leu Val Glu Leu Gln Glu
    130                 135                 140

Leu Cys Arg Arg Arg His Asp Glu Asp Val Pro Asp Phe Leu Met His
145                 150                 155                 160

Lys Leu Thr Ala Ser Asp Thr Gly Lys Thr Cys Leu Met Lys Ala Leu
                165                 170                 175

Leu Asn Ile Asn Pro Asn Thr Lys Glu Ile Val Arg Ile Leu Leu Ala
            180                 185                 190

Phe Ala Glu Glu Asn Asp Ile Leu Gly Arg Phe Ile Asn Ala Glu Tyr
        195                 200                 205

Thr Glu Glu Ala Tyr Glu Gly Gln Thr Ala Leu Asn Ile Ala Ile Glu
    210                 215                 220

Arg Arg Gln Gly Asp Ile Ala Ala Leu Leu Ile Ala Ala Gly Ala Asp
225                 230                 235                 240

Val Asn Ala His Ala Lys Gly Ala Phe Phe Asn Pro Lys Tyr Gln His
                245                 250                 255

Glu Gly Phe Tyr Phe Gly Glu Thr Pro Leu Ala Leu Ala Ala Cys Thr
            260                 265                 270

Asn Gln Pro Glu Ile Val Gln Leu Leu Met Glu His Glu Gln Thr Asp
        275                 280                 285

Ile Thr Ser Arg Asp Ser Arg Gly Asn Asn Ile Leu His Ala Leu Val
    290                 295                 300

Thr Val Ala Glu Asp Phe Lys Thr Gln Asn Asp Phe Val Lys Arg Met
305                 310                 315                 320

Tyr Asp Met Ile Leu Leu Arg Ser Gly Asn Trp Glu Leu Glu Thr Thr
                325                 330                 335

Arg Asn Asn Asp Gly Leu Thr Pro Leu Gln Leu Ala Ala Lys Met Gly
            340                 345                 350
```

-continued

```
Lys Ala Glu Ile Leu Lys Tyr Ile Leu Ser Arg Glu Ile Lys Glu Lys
        355                 360                 365

Arg Leu Arg Ser Leu Ser Arg Lys Phe Thr Asp Trp Ala Tyr Gly Pro
    370                 375                 380

Val Ser Ser Ser Leu Tyr Asp Leu Thr Asn Val Asp Thr Thr Thr Asp
385                 390                 395                 400

Asn Ser Val Leu Glu Ile Thr Val Tyr Asn Thr Asn Ile Asp Asn Arg
                405                 410                 415

His Glu Met Leu Thr Leu Glu Pro Leu His Thr Leu Leu His Met Lys
            420                 425                 430

Trp Lys Lys Phe Ala Lys His Met Phe Phe Leu Ser Phe Cys Phe Tyr
        435                 440                 445

Phe Phe Tyr Asn Ile Thr Leu Thr Leu Val Ser Tyr Tyr Arg Pro Arg
    450                 455                 460

Glu Glu Glu Ala Ile Pro His Pro Leu Ala Leu Thr His Lys Met Gly
465                 470                 475                 480

Trp Leu Gln Leu Leu Gly Arg Met Phe Val Leu Ile Trp Ala Met Cys
                485                 490                 495

Ile Ser Val Lys Glu Gly Ile Ala Ile Phe Leu Leu Arg Pro Ser Asp
            500                 505                 510

Leu Gln Ser Ile Leu Ser Asp Ala Trp Phe His Phe Val Phe Phe Ile
        515                 520                 525

Gln Ala Val Leu Val Ile Leu Ser Val Phe Leu Tyr Leu Phe Ala Tyr
    530                 535                 540

Lys Glu Tyr Leu Ala Cys Leu Val Leu Ala Met Ala Leu Gly Trp Ala
545                 550                 555                 560

Asn Met Leu Tyr Tyr Thr Arg Gly Phe Gln Ser Met Gly Met Tyr Ser
                565                 570                 575

Val Met Ile Gln Lys Val Ile Leu His Asp Val Leu Lys Phe Leu Phe
            580                 585                 590

Val Tyr Ile Val Phe Leu Leu Gly Phe Gly Val Ala Leu Ala Ser Leu
        595                 600                 605

Ile Glu Lys Cys Pro Lys Asp Asn Lys Asp Cys Ser Ser Tyr Gly Ser
    610                 615                 620

Phe Ser Asp Ala Val Leu Glu Leu Phe Lys Leu Thr Ile Gly Leu Gly
625                 630                 635                 640

Asp Leu Asn Ile Gln Gln Asn Ser Lys Tyr Pro Ile Leu Phe Leu Phe
                645                 650                 655

Leu Leu Ile Thr Tyr Val Ile Leu Thr Phe Val Leu Leu Leu Asn Met
            660                 665                 670

Leu Ile Ala Leu Met Gly Glu Thr Val Glu Asn Val Ser Lys Glu Ser
        675                 680                 685

Glu Arg Ile Trp Arg Leu Gln Arg Ala Arg Thr Ile Leu Glu Phe Glu
    690                 695                 700

Lys Met Leu Pro Glu Trp Leu Arg Ser Arg Phe Arg Met Gly Glu Leu
705                 710                 715                 720

Cys Lys Val Ala Glu Asp Asp Phe Arg Leu Cys Leu Arg Ile Asn Glu
                725                 730                 735

Val Lys Trp Thr Glu Trp Lys Thr His Val Ser Phe Leu Asn Glu Asp
            740                 745                 750

Pro Gly Pro Val Arg Arg Thr Asp Phe Asn Lys Ile Gln Asp Ser Ser
```

-continued

```
        755                 760                 765

Arg Asn Asn Ser Lys Thr Thr Leu Asn Ala Phe Glu Glu Val Glu Glu
    770                 775                 780

Phe Pro Glu Thr Ser Val
785                 790


<210> SEQ ID NO 23
<211> LENGTH: 2373
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(2373)

<400> SEQUENCE: 23

atg aaa gcc cac ccc aag gag atg gtg cct ctc atg ggc aag aga gtt       48
Met Lys Ala His Pro Lys Glu Met Val Pro Leu Met Gly Lys Arg Val
 1               5                  10                  15

gct gcc ccc agt ggg aac cct gcc gtc ctg cca gag aag agg ccg gcg       96
Ala Ala Pro Ser Gly Asn Pro Ala Val Leu Pro Glu Lys Arg Pro Ala
            20                  25                  30

gag atc acc ccc aca aag aag agt gca cac ttc ttc ctg gag ata gaa      144
Glu Ile Thr Pro Thr Lys Lys Ser Ala His Phe Phe Leu Glu Ile Glu
        35                  40                  45

ggg ttt gaa ccc aac ccc aca gtt gcc aag acc tct cct cct gtc ttc      192
Gly Phe Glu Pro Asn Pro Thr Val Ala Lys Thr Ser Pro Pro Val Phe
    50                  55                  60

tcc aag ccc atg gat tcc aac atc cgg cag tgc atc tct ggt aac tgt      240
Ser Lys Pro Met Asp Ser Asn Ile Arg Gln Cys Ile Ser Gly Asn Cys
 65                  70                  75                  80

gat gac atg gac tcc ccc cag tct cct caa gat gat gtg aca gag acc      288
Asp Asp Met Asp Ser Pro Gln Ser Pro Gln Asp Asp Val Thr Glu Thr
                85                  90                  95

cca tcc aat ccc aac agc ccc agt gca cag ctg gcc aag gaa gag cag      336
Pro Ser Asn Pro Asn Ser Pro Ser Ala Gln Leu Ala Lys Glu Glu Gln
            100                 105                 110

agg agg aaa aag agg cgg ctg aag aag cgc atc ttt gca gcc gtg tct      384
Arg Arg Lys Lys Arg Arg Leu Lys Lys Arg Ile Phe Ala Ala Val Ser
        115                 120                 125

gag ggc tgc gtg gag gag ttg gta gag ttg ctg gtg gag ctg cag gag      432
Glu Gly Cys Val Glu Glu Leu Val Glu Leu Leu Val Glu Leu Gln Glu
    130                 135                 140

ctt tgc agg cgg cgc cat gat gag gat gtg cct gac ttc ctc atg cac      480
Leu Cys Arg Arg Arg His Asp Glu Asp Val Pro Asp Phe Leu Met His
145                 150                 155                 160

aag ctg acg gcc tcc gac acg ggg aag acc tgc ctg atg aag gcc ttg      528
Lys Leu Thr Ala Ser Asp Thr Gly Lys Thr Cys Leu Met Lys Ala Leu
                165                 170                 175

tta aac atc aac ccc aac acc aag gag atc gtg cgg atc ctg ctt gcc      576
Leu Asn Ile Asn Pro Asn Thr Lys Glu Ile Val Arg Ile Leu Leu Ala
            180                 185                 190

ttt gct gaa gag aac gac atc ctg ggc agg ttc atc aac gcc gag tac      624
Phe Ala Glu Glu Asn Asp Ile Leu Gly Arg Phe Ile Asn Ala Glu Tyr
        195                 200                 205

aca gag gag gcc tat gaa ggg cag acg gcg ctg aac atc gcc atc gag      672
Thr Glu Glu Ala Tyr Glu Gly Gln Thr Ala Leu Asn Ile Ala Ile Glu
    210                 215                 220

cgg cgg cag ggg gac atc gca gcc ctg ctc atc gcc gcc ggc gcc gac      720
```

-continued

```
Arg Arg Gln Gly Asp Ile Ala Ala Leu Leu Ile Ala Ala Gly Ala Asp
225                 230                 235                 240

gtc aac gcg cac gcc aag ggg gcc ttc ttc aac ccc aag tac caa cac      768
Val Asn Ala His Ala Lys Gly Ala Phe Phe Asn Pro Lys Tyr Gln His
                245                 250                 255

gaa ggc ttc tac ttc ggt gag acg ccc ctg gcc ctg gca gca tgc acc      816
Glu Gly Phe Tyr Phe Gly Glu Thr Pro Leu Ala Leu Ala Ala Cys Thr
            260                 265                 270

aac cag ccc gag att gtg cag ctg ctg atg gag cac gag cag acg gac      864
Asn Gln Pro Glu Ile Val Gln Leu Leu Met Glu His Glu Gln Thr Asp
        275                 280                 285

atc acc tcg cgg gac tca cga ggc aac aac atc ctt cac gcc ctg gtg      912
Ile Thr Ser Arg Asp Ser Arg Gly Asn Asn Ile Leu His Ala Leu Val
    290                 295                 300

acc gtg gcc gag gac ttc aag acg cag aat gac ttt gtg aag cgc atg      960
Thr Val Ala Glu Asp Phe Lys Thr Gln Asn Asp Phe Val Lys Arg Met
305                 310                 315                 320

tac gac atg atc cta ctg cgg agt ggc aac tgg gag ctg gag acc act     1008
Tyr Asp Met Ile Leu Leu Arg Ser Gly Asn Trp Glu Leu Glu Thr Thr
                325                 330                 335

cgc aac aac gat ggc ctc acg ccg ctg cag ctg gcc gcc aag atg ggc     1056
Arg Asn Asn Asp Gly Leu Thr Pro Leu Gln Leu Ala Ala Lys Met Gly
            340                 345                 350

aag gcg gag atc ctg aag tac atc ctc agt cgt gag atc aag gag aag     1104
Lys Ala Glu Ile Leu Lys Tyr Ile Leu Ser Arg Glu Ile Lys Glu Lys
        355                 360                 365

cgg ctc cgg agc ctg tcc agg aag ttc acc gac tgg gcg tac gga ccc     1152
Arg Leu Arg Ser Leu Ser Arg Lys Phe Thr Asp Trp Ala Tyr Gly Pro
    370                 375                 380

gtg tca tcc tcc ctc tac gac ctc acc aac gtg gac acc acc acg gac     1200
Val Ser Ser Ser Leu Tyr Asp Leu Thr Asn Val Asp Thr Thr Thr Asp
385                 390                 395                 400

aac tca gtg ctg gaa atc act gtc tac aac acc aac atc gac aac cgg     1248
Asn Ser Val Leu Glu Ile Thr Val Tyr Asn Thr Asn Ile Asp Asn Arg
                405                 410                 415

cat gag atg ctg acc ctg gag ccg ctg cac acg ctg ctg cat atg aag     1296
His Glu Met Leu Thr Leu Glu Pro Leu His Thr Leu Leu His Met Lys
            420                 425                 430

tgg aag aag ttt gcc aag cac atg ttc ttt ctg tcc ttc tgc ttt tat     1344
Trp Lys Lys Phe Ala Lys His Met Phe Phe Leu Ser Phe Cys Phe Tyr
        435                 440                 445

ttc ttc tac aac atc acc ctg acc ctc gtc tcg tac tac cgc ccc cgg     1392
Phe Phe Tyr Asn Ile Thr Leu Thr Leu Val Ser Tyr Tyr Arg Pro Arg
    450                 455                 460

gag gag gag gcc atc ccg cac ccc ttg gcc ctg acg cac aag atg ggg     1440
Glu Glu Glu Ala Ile Pro His Pro Leu Ala Leu Thr His Lys Met Gly
465                 470                 475                 480

tgg ctg cag ctc cta ggg agg atg ttt gtg ctc atc tgg gcc atg tgc     1488
Trp Leu Gln Leu Leu Gly Arg Met Phe Val Leu Ile Trp Ala Met Cys
                485                 490                 495

atc tct gtg aaa gag ggc att gcc atc ttc ctg ctg aga ccc tcg gat     1536
Ile Ser Val Lys Glu Gly Ile Ala Ile Phe Leu Leu Arg Pro Ser Asp
            500                 505                 510

ctg cag tcc atc ctc tcg gat gcc tgg ttc cac ttt gtc ttt ttt atc     1584
Leu Gln Ser Ile Leu Ser Asp Ala Trp Phe His Phe Val Phe Phe Ile
        515                 520                 525

caa gct gtg ctt gtg ata ctg tct gtc ttc ttg tac ttg ttt gcc tac     1632
```

-continued

```
Gln Ala Val Leu Val Ile Leu Ser Val Phe Leu Tyr Leu Phe Ala Tyr
    530                 535                 540

aaa gag tac ctc gcc tgc ctc gtg ctg gcc atg gcc ctg ggc tgg gcg     1680
Lys Glu Tyr Leu Ala Cys Leu Val Leu Ala Met Ala Leu Gly Trp Ala
545                 550                 555                 560

aac atg ctc tac tat acg cgg ggt ttc cag tcc atg ggc atg tac agc     1728
Asn Met Leu Tyr Tyr Thr Arg Gly Phe Gln Ser Met Gly Met Tyr Ser
                565                 570                 575

gtc atg atc cag aag gtc att ttg cat gat gtt ctg aag ttc ttg ttt     1776
Val Met Ile Gln Lys Val Ile Leu His Asp Val Leu Lys Phe Leu Phe
            580                 585                 590

gta tat atc gtg ttt ttg ctt gga ttt gga gta gcc ttg gcc tcg ctg     1824
Val Tyr Ile Val Phe Leu Leu Gly Phe Gly Val Ala Leu Ala Ser Leu
        595                 600                 605

atc gag aag tgt ccc aaa gac aac aag gac tgc agc tcc tac ggc agc     1872
Ile Glu Lys Cys Pro Lys Asp Asn Lys Asp Cys Ser Ser Tyr Gly Ser
    610                 615                 620

ttc agt gac gca gtg ctg gaa ctc ttc aag ctc acc ata ggc ctg ggt     1920
Phe Ser Asp Ala Val Leu Glu Leu Phe Lys Leu Thr Ile Gly Leu Gly
625                 630                 635                 640

gat ctg aac atc cag cag aac tcc aag tat ccc att ctc ttt ctg ttc     1968
Asp Leu Asn Ile Gln Gln Asn Ser Lys Tyr Pro Ile Leu Phe Leu Phe
                645                 650                 655

ctg ctc atc acc tat gtc atc ctc acc ttt gtt ctc ctc ctc aac atg     2016
Leu Leu Ile Thr Tyr Val Ile Leu Thr Phe Val Leu Leu Leu Asn Met
            660                 665                 670

ctc att gct ctg atg ggc gag act gtg gag aac gtc tcc aag gag agc     2064
Leu Ile Ala Leu Met Gly Glu Thr Val Glu Asn Val Ser Lys Glu Ser
        675                 680                 685

gaa cgc atc tgg cgc ctg cag aga gcc agg acc atc ttg gag ttt gag     2112
Glu Arg Ile Trp Arg Leu Gln Arg Ala Arg Thr Ile Leu Glu Phe Glu
    690                 695                 700

aaa atg tta cca gaa tgg ctg agg agc aga ttc cgg atg gga gag ctg     2160
Lys Met Leu Pro Glu Trp Leu Arg Ser Arg Phe Arg Met Gly Glu Leu
705                 710                 715                 720

tgc aaa gtg gcc gag gat gat ttc cga ctg tgt ttg cgg atc aat gag     2208
Cys Lys Val Ala Glu Asp Asp Phe Arg Leu Cys Leu Arg Ile Asn Glu
                725                 730                 735

gtg aag tgg act gaa tgg aag acg cac gtc tcc ttc ctt aac gaa gac     2256
Val Lys Trp Thr Glu Trp Lys Thr His Val Ser Phe Leu Asn Glu Asp
            740                 745                 750

ccg ggg cct gta aga cga aca gat ttc aac aaa atc caa gat tct tcc     2304
Pro Gly Pro Val Arg Arg Thr Asp Phe Asn Lys Ile Gln Asp Ser Ser
        755                 760                 765

agg aac aac agc aaa acc act ctc aat gca ttt gaa gaa gtc gag gaa     2352
Arg Asn Asn Ser Lys Thr Thr Leu Asn Ala Phe Glu Glu Val Glu Glu
    770                 775                 780

ttc ccg gaa acc tcg gtg tag                                         2373
Phe Pro Glu Thr Ser Val  *
785                 790
```

<210> SEQ ID NO 24
<211> LENGTH: 3042
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(3042)

-continued

<400> SEQUENCE: 24

```
atg gtt gga gga tgc agg tgg aca gaa gac gtg gag cct gca gaa gta       48
Met Val Gly Gly Cys Arg Trp Thr Glu Asp Val Glu Pro Ala Glu Val
 1               5                  10                  15

aag gaa aag atg tcc ttt cgg gca gcc agg ctc agc atg agg aac aga       96
Lys Glu Lys Met Ser Phe Arg Ala Ala Arg Leu Ser Met Arg Asn Arg
            20                  25                  30

agg aat gac act ctg gac agc acc cgg acc ctg tac tcc agc gcg tct      144
Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu Tyr Ser Ser Ala Ser
        35                  40                  45

cgg agc aca gac ttg tct tac agt gaa agc gac ttg gtg aat ttt att      192
Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser Asp Leu Val Asn Phe Ile
    50                  55                  60

caa gca aat ttt aag aaa cga gaa tgt gtc ttc ttt acc aaa gat tcc      240
Gln Ala Asn Phe Lys Lys Arg Glu Cys Val Phe Phe Thr Lys Asp Ser
65                  70                  75                  80

aag gcc acg gag aat gtg tgc aag tgt ggc tat gcc cag agc cag cac      288
Lys Ala Thr Glu Asn Val Cys Lys Cys Gly Tyr Ala Gln Ser Gln His
                85                  90                  95

atg gaa ggc acc cag atc aac caa agt gag aaa tgg aac tac aag aaa      336
Met Glu Gly Thr Gln Ile Asn Gln Ser Glu Lys Trp Asn Tyr Lys Lys
            100                 105                 110

cac acc aag gaa ttt cct acc gac gcc ttt ggg gat att cag ttt gag      384
His Thr Lys Glu Phe Pro Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu
        115                 120                 125

aca ctg ggg aag aaa ggg aag tat ata cgt ctg tcc tgc gac acg gac      432
Thr Leu Gly Lys Lys Gly Lys Tyr Ile Arg Leu Ser Cys Asp Thr Asp
    130                 135                 140

gcg gaa atc ctt tac gag ctg ctg acc cag cac tgg cac ctg aaa aca      480
Ala Glu Ile Leu Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr
145                 150                 155                 160

ccc aac ctg gtc att tct gtg acc ggg ggc gcc aag aac ttc gcc ctg      528
Pro Asn Leu Val Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Ala Leu
                165                 170                 175

aag ccg cgc atg cgc aag atc ttc agc cgg ctc atc tac atc gcg cag      576
Lys Pro Arg Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln
            180                 185                 190

tcc aaa ggt gct tgg att ctc acg gga ggc acc cat tat ggc ctg atg      624
Ser Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His Tyr Gly Leu Met
        195                 200                 205

aag tac ctc ggg gag gtg gtg aga gat aac acc atc agc agg agt tca      672
Lys Tyr Leu Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser Ser
    210                 215                 220

gag gag aat att gtg gcc att ggc ata gca gct tgg ggc atg gtc tcc      720
Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Ala Trp Gly Met Val Ser
225                 230                 235                 240

aac cgg gac acc ctc atc agg aat tgc gat gct gag ggc tat ttt tta      768
Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly Tyr Phe Leu
                245                 250                 255

gcc cag tac ctt atg gat gac ttc aca aga gat cca ctg tat atc ctg      816
Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro Leu Tyr Ile Leu
            260                 265                 270

gac aac aac cac aca cat ttg ctg ctc gtg gac aat ggc tgt cat gga      864
Asp Asn Asn His Thr His Leu Leu Leu Val Asp Asn Gly Cys His Gly
        275                 280                 285

cat ccc act gtc gaa gca aag ctc cgg aat cag cta gag aag tat atc      912
His Pro Thr Val Glu Ala Lys Leu Arg Asn Gln Leu Glu Lys Tyr Ile
```

-continued

```
    290                 295                 300

tct gag cgc act att caa gat tcc aac tat ggt ggc aag atc ccc att      960
Ser Glu Arg Thr Ile Gln Asp Ser Asn Tyr Gly Gly Lys Ile Pro Ile
305                 310                 315                 320

gtg tgt ttt gcc caa gga ggt gga aaa gag act ttg aaa gcc atc aat     1008
Val Cys Phe Ala Gln Gly Gly Gly Lys Glu Thr Leu Lys Ala Ile Asn
                325                 330                 335

acc tcc atc aaa aat aaa att cct tgt gtg gtg gtg gaa ggc tcg ggc     1056
Thr Ser Ile Lys Asn Lys Ile Pro Cys Val Val Val Glu Gly Ser Gly
            340                 345                 350

cag atc gct gat gtg atc gct agc ctg gtg gag gtg gag gat gcc ctg     1104
Gln Ile Ala Asp Val Ile Ala Ser Leu Val Glu Val Glu Asp Ala Leu
        355                 360                 365

aca tct tct gcc gtc aag gag aag ctg gtg cgc ttt tta ccc cgc acg     1152
Thr Ser Ser Ala Val Lys Glu Lys Leu Val Arg Phe Leu Pro Arg Thr
    370                 375                 380

gtg tcc cgg ctg cct gag gag gag act gag agt tgg atc aaa tgg ctc     1200
Val Ser Arg Leu Pro Glu Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu
385                 390                 395                 400

aaa gaa att ctc gaa tgt tct cac cta tta aca gtt att aaa atg gaa     1248
Lys Glu Ile Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu
                405                 410                 415

gaa gct ggg gat gaa att gtg agc aat gcc atc tcc tac gct cta tac     1296
Glu Ala Gly Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr
            420                 425                 430

aaa gcc ttc agc acc agt gag caa gac aag gat aac tgg aat ggg cag     1344
Lys Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln
        435                 440                 445

ctg aag ctt ctg ctg gag tgg aac cag ctg gac tta gcc aat gat gag     1392
Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp Glu
    450                 455                 460

att ttc acc aat gac cgc cga tgg gag aag agc aaa ccg agg ctc aga     1440
Ile Phe Thr Asn Asp Arg Arg Trp Glu Lys Ser Lys Pro Arg Leu Arg
465                 470                 475                 480

gac aca ata atc cag gtc aca tgg ctg gaa aat ggt aga atc aag gtt     1488
Asp Thr Ile Ile Gln Val Thr Trp Leu Glu Asn Gly Arg Ile Lys Val
                485                 490                 495

gag agc aaa gat gtg act gac ggc aaa gcc tct tct cat atg ctg gtg     1536
Glu Ser Lys Asp Val Thr Asp Gly Lys Ala Ser Ser His Met Leu Val
            500                 505                 510

gtt ctc aag tct gct gac ctt caa gaa gtc atg ttt acg gct ctc ata     1584
Val Leu Lys Ser Ala Asp Leu Gln Glu Val Met Phe Thr Ala Leu Ile
        515                 520                 525

aag gac aga ccc aag ttt gtc cgc ctc ttt ctg gag aat ggc ttg aac     1632
Lys Asp Arg Pro Lys Phe Val Arg Leu Phe Leu Glu Asn Gly Leu Asn
    530                 535                 540

cta cgg aag ttt ctc acc cat gat gtc ctc act gaa ctc ttc tcc aac     1680
Leu Arg Lys Phe Leu Thr His Asp Val Leu Thr Glu Leu Phe Ser Asn
545                 550                 555                 560

cac ttc agc acg ctt gtg tac cgg aat ctg cag atc gcc aag aat tcc     1728
His Phe Ser Thr Leu Val Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser
                565                 570                 575

tat aat gat gcc ctc ctc acg ttt gtc tgg aaa ctg gtt gcg aac ttc     1776
Tyr Asn Asp Ala Leu Leu Thr Phe Val Trp Lys Leu Val Ala Asn Phe
            580                 585                 590

cga aga ggc ttc cgg aag gaa gac aga aat ggc cgg gac gag atg gac     1824
Arg Arg Gly Phe Arg Lys Glu Asp Arg Asn Gly Arg Asp Glu Met Asp
```

-continued

```
        595                 600                 605

ata gaa ctc cac gac gtg tct cct att act cgg cac ccc ctg caa gct     1872
Ile Glu Leu His Asp Val Ser Pro Ile Thr Arg His Pro Leu Gln Ala
    610                 615                 620

ctc ttc atc tgg gcc att ctt cag aat aag aag gaa ctc tcc aaa gtc     1920
Leu Phe Ile Trp Ala Ile Leu Gln Asn Lys Lys Glu Leu Ser Lys Val
625                 630                 635                 640

att tgg gag cag acc agg ggc tgc act ctg gca gcc ctg gga gcc agc     1968
Ile Trp Glu Gln Thr Arg Gly Cys Thr Leu Ala Ala Leu Gly Ala Ser
                645                 650                 655

aag ctt ctg aag act ctg gcc aaa gtg aag aac gac atc aat gct gct     2016
Lys Leu Leu Lys Thr Leu Ala Lys Val Lys Asn Asp Ile Asn Ala Ala
            660                 665                 670

ggg gag tcc gag gag ctg gct aat gag tac ctg acc cgg gct gtt ggt     2064
Gly Glu Ser Glu Glu Leu Ala Asn Glu Tyr Leu Thr Arg Ala Val Gly
        675                 680                 685

gag tcc aca gtg tgg aat gct gtg gtg ggc gcg gat ctg cca tgt ggc     2112
Glu Ser Thr Val Trp Asn Ala Val Val Gly Ala Asp Leu Pro Cys Gly
    690                 695                 700

aca gac att gcc agc ggc act cat aga cca gat ggt gga gag ctg ttc     2160
Thr Asp Ile Ala Ser Gly Thr His Arg Pro Asp Gly Gly Glu Leu Phe
705                 710                 715                 720

act gag tgt tac agc agc gat gaa gac ttg gca gaa cag ctg ctg gtc     2208
Thr Glu Cys Tyr Ser Ser Asp Glu Asp Leu Ala Glu Gln Leu Leu Val
                725                 730                 735

tat tcc tgt gaa gct tgg ggt gga agc aac tgt ctg gag ctg gcg gtg     2256
Tyr Ser Cys Glu Ala Trp Gly Gly Ser Asn Cys Leu Glu Leu Ala Val
            740                 745                 750

gag gcc aca gac cag cat ttc atc gcc cag cct ggg gtc cag aat ttt     2304
Glu Ala Thr Asp Gln His Phe Ile Ala Gln Pro Gly Val Gln Asn Phe
        755                 760                 765

ctt tct aag caa tgg tat gga gag att tcc cga gac acc aag aac tgg     2352
Leu Ser Lys Gln Trp Tyr Gly Glu Ile Ser Arg Asp Thr Lys Asn Trp
    770                 775                 780

aag att atc ctg tgt ctg ttt att ata ccc ttg gtg ggc tgt ggc ttt     2400
Lys Ile Ile Leu Cys Leu Phe Ile Ile Pro Leu Val Gly Cys Gly Phe
785                 790                 795                 800

gta tca ttt agg aag aaa cct gtc gac aag cac aag aag ctg ctt tgg     2448
Val Ser Phe Arg Lys Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp
                805                 810                 815

tac tat gtg gcg ttc ttc acc tcc ccc ttc gtg gtc ttc tcc tgg aat     2496
Tyr Tyr Val Ala Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn
            820                 825                 830

gtg gtc ttc tac atc gcc ttc ctc ctg ctg ttt gcc tac gtg ctg ctc     2544
Val Val Phe Tyr Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu
        835                 840                 845

atg gat ttc cat tcg gtg cca cac ccc ccc gag ctg gtc ctg tac tcg     2592
Met Asp Phe His Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser
    850                 855                 860

ctg gtc ttt gtc ctc ttc tgt gat gaa gtg aga cag ggc cgg ccg gct     2640
Leu Val Phe Val Leu Phe Cys Asp Glu Val Arg Gln Gly Arg Pro Ala
865                 870                 875                 880

gct ccc agt gcg ggg ccc gcc aag ccc acg ccc acc cgg aac tcc atc     2688
Ala Pro Ser Ala Gly Pro Ala Lys Pro Thr Pro Thr Arg Asn Ser Ile
                885                 890                 895

tgg ccc gca agc tcc aca cgc agc ccc ggt tcc cgc tca cgc cac tcc     2736
Trp Pro Ala Ser Ser Thr Arg Ser Pro Gly Ser Arg Ser Arg His Ser
```

-continued

```
            900                 905                 910

ttc cac act tcc ctg caa gct gag ggt gcc agc tct ggc ctt ggc cag     2784
Phe His Thr Ser Leu Gln Ala Glu Gly Ala Ser Ser Gly Leu Gly Gln
        915                 920                 925

ccc aga aag ggg ctc cca cag tgc agc ggt ggg ctg aag ggc tcc tca     2832
Pro Arg Lys Gly Leu Pro Gln Cys Ser Gly Gly Leu Lys Gly Ser Ser
    930                 935                 940

agt gcc gcc aaa gtg gga gcc cag gca gag gag gtg ccg aga gca agc     2880
Ser Ala Ala Lys Val Gly Ala Gln Ala Glu Glu Val Pro Arg Ala Ser
945                 950                 955                 960

gag ggc tgt gag gac tgc cag cac gct gtc acc tct cag aag cgt aag     2928
Glu Gly Cys Glu Asp Cys Gln His Ala Val Thr Ser Gln Lys Arg Lys
                965                 970                 975

aca gca atg gac caa aca gac gaa gat ctc ttc ccc tat gga gca ttc     2976
Thr Ala Met Asp Gln Thr Asp Glu Asp Leu Phe Pro Tyr Gly Ala Phe
            980                 985                 990

tac cag ttc ctg atg att tcc agg agc ttt cga gga gag gag atg agc     3024
Tyr Gln Phe Leu Met Ile Ser Arg Ser Phe Arg Gly Glu Glu Met Ser
         995                 1000                1005

atc ggc aag cag cac taa                                             3042
Ile Gly Lys Gln His  *
    1010


<210> SEQ ID NO 25
<211> LENGTH: 1013
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 25

Met Val Gly Gly Cys Arg Trp Thr Glu Asp Val Glu Pro Ala Glu Val
 1               5                  10                  15

Lys Glu Lys Met Ser Phe Arg Ala Ala Arg Leu Ser Met Arg Asn Arg
            20                  25                  30

Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu Tyr Ser Ser Ala Ser
        35                  40                  45

Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser Asp Leu Val Asn Phe Ile
    50                  55                  60

Gln Ala Asn Phe Lys Lys Arg Glu Cys Val Phe Phe Thr Lys Asp Ser
65                  70                  75                  80

Lys Ala Thr Glu Asn Val Cys Lys Cys Gly Tyr Ala Gln Ser Gln His
                85                  90                  95

Met Glu Gly Thr Gln Ile Asn Gln Ser Glu Lys Trp Asn Tyr Lys Lys
            100                 105                 110

His Thr Lys Glu Phe Pro Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu
        115                 120                 125

Thr Leu Gly Lys Lys Gly Lys Tyr Ile Arg Leu Ser Cys Asp Thr Asp
    130                 135                 140

Ala Glu Ile Leu Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr
145                 150                 155                 160

Pro Asn Leu Val Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Ala Leu
                165                 170                 175

Lys Pro Arg Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln
            180                 185                 190

Ser Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His Tyr Gly Leu Met
        195                 200                 205
```

-continued

```
Lys Tyr Leu Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser Ser
    210                 215                 220

Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Ala Trp Gly Met Val Ser
225                 230                 235                 240

Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly Tyr Phe Leu
                245                 250                 255

Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro Leu Tyr Ile Leu
            260                 265                 270

Asp Asn Asn His Thr His Leu Leu Leu Val Asp Asn Gly Cys His Gly
        275                 280                 285

His Pro Thr Val Glu Ala Lys Leu Arg Asn Gln Leu Glu Lys Tyr Ile
    290                 295                 300

Ser Glu Arg Thr Ile Gln Asp Ser Asn Tyr Gly Gly Lys Ile Pro Ile
305                 310                 315                 320

Val Cys Phe Ala Gln Gly Gly Gly Lys Glu Thr Leu Lys Ala Ile Asn
                325                 330                 335

Thr Ser Ile Lys Asn Lys Ile Pro Cys Val Val Val Glu Gly Ser Gly
            340                 345                 350

Gln Ile Ala Asp Val Ile Ala Ser Leu Val Glu Val Glu Asp Ala Leu
        355                 360                 365

Thr Ser Ser Ala Val Lys Glu Lys Leu Val Arg Phe Leu Pro Arg Thr
    370                 375                 380

Val Ser Arg Leu Pro Glu Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu
385                 390                 395                 400

Lys Glu Ile Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu
                405                 410                 415

Glu Ala Gly Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr
            420                 425                 430

Lys Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln
        435                 440                 445

Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp Glu
    450                 455                 460

Ile Phe Thr Asn Asp Arg Arg Trp Glu Lys Ser Lys Pro Arg Leu Arg
465                 470                 475                 480

Asp Thr Ile Ile Gln Val Thr Trp Leu Glu Asn Gly Arg Ile Lys Val
                485                 490                 495

Glu Ser Lys Asp Val Thr Asp Gly Lys Ala Ser Ser His Met Leu Val
            500                 505                 510

Val Leu Lys Ser Ala Asp Leu Gln Glu Val Met Phe Thr Ala Leu Ile
        515                 520                 525

Lys Asp Arg Pro Lys Phe Val Arg Leu Phe Leu Glu Asn Gly Leu Asn
    530                 535                 540

Leu Arg Lys Phe Leu Thr His Asp Val Leu Thr Glu Leu Phe Ser Asn
545                 550                 555                 560

His Phe Ser Thr Leu Val Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser
                565                 570                 575

Tyr Asn Asp Ala Leu Leu Thr Phe Val Trp Lys Leu Val Ala Asn Phe
            580                 585                 590

Arg Arg Gly Phe Arg Lys Glu Asp Arg Asn Gly Arg Asp Glu Met Asp
        595                 600                 605
```

-continued

```
Ile Glu Leu His Asp Val Ser Pro Ile Thr Arg His Pro Leu Gln Ala
    610                 615                 620

Leu Phe Ile Trp Ala Ile Leu Gln Asn Lys Lys Glu Leu Ser Lys Val
625                 630                 635                 640

Ile Trp Glu Gln Thr Arg Gly Cys Thr Leu Ala Ala Leu Gly Ala Ser
                645                 650                 655

Lys Leu Leu Lys Thr Leu Ala Lys Val Lys Asn Asp Ile Asn Ala Ala
            660                 665                 670

Gly Glu Ser Glu Glu Leu Ala Asn Glu Tyr Leu Thr Arg Ala Val Gly
        675                 680                 685

Glu Ser Thr Val Trp Asn Ala Val Val Gly Ala Asp Leu Pro Cys Gly
    690                 695                 700

Thr Asp Ile Ala Ser Gly Thr His Arg Pro Asp Gly Gly Glu Leu Phe
705                 710                 715                 720

Thr Glu Cys Tyr Ser Ser Asp Glu Asp Leu Ala Glu Gln Leu Leu Val
                725                 730                 735

Tyr Ser Cys Glu Ala Trp Gly Gly Ser Asn Cys Leu Glu Leu Ala Val
            740                 745                 750

Glu Ala Thr Asp Gln His Phe Ile Ala Gln Pro Gly Val Gln Asn Phe
        755                 760                 765

Leu Ser Lys Gln Trp Tyr Gly Glu Ile Ser Arg Asp Thr Lys Asn Trp
    770                 775                 780

Lys Ile Ile Leu Cys Leu Phe Ile Ile Pro Leu Val Gly Cys Gly Phe
785                 790                 795                 800

Val Ser Phe Arg Lys Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp
                805                 810                 815

Tyr Tyr Val Ala Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn
            820                 825                 830

Val Val Phe Tyr Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu
        835                 840                 845

Met Asp Phe His Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser
    850                 855                 860

Leu Val Phe Val Leu Phe Cys Asp Glu Val Arg Gln Gly Arg Pro Ala
865                 870                 875                 880

Ala Pro Ser Ala Gly Pro Ala Lys Pro Thr Pro Thr Arg Asn Ser Ile
                885                 890                 895

Trp Pro Ala Ser Ser Thr Arg Ser Pro Gly Ser Arg Ser Arg His Ser
            900                 905                 910

Phe His Thr Ser Leu Gln Ala Glu Gly Ala Ser Ser Gly Leu Gly Gln
        915                 920                 925

Pro Arg Lys Gly Leu Pro Gln Cys Ser Gly Gly Leu Lys Gly Ser Ser
    930                 935                 940

Ser Ala Ala Lys Val Gly Ala Gln Ala Glu Glu Val Pro Arg Ala Ser
945                 950                 955                 960

Glu Gly Cys Glu Asp Cys Gln His Ala Val Thr Ser Gln Lys Arg Lys
                965                 970                 975

Thr Ala Met Asp Gln Thr Asp Glu Asp Leu Phe Pro Tyr Gly Ala Phe
            980                 985                 990

Tyr Gln Phe Leu Met Ile Ser Arg Ser Phe Arg Gly Glu Glu Met Ser
        995                 1000                1005

Ile Gly Lys Gln His
```

-continued

```
    1010


<210> SEQ ID NO 26
<211> LENGTH: 3042
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(3042)

<400> SEQUENCE: 26

atg gtt gga gga tgc agg tgg aca gaa gac gtg gag cct gca gaa gta       48
Met Val Gly Gly Cys Arg Trp Thr Glu Asp Val Glu Pro Ala Glu Val
 1               5                  10                  15

aag gaa aag atg tcc ttt cgg gca gcc agg ctc agc atg agg aac aga       96
Lys Glu Lys Met Ser Phe Arg Ala Ala Arg Leu Ser Met Arg Asn Arg
            20                  25                  30

agg aat gac act ctg gac agc acc cgg acc ctg tac tcc agc gcg tct      144
Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu Tyr Ser Ser Ala Ser
        35                  40                  45

cgg agc aca gac ttg tct tac agt gaa agc gac ttg gtg aat ttt att      192
Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser Asp Leu Val Asn Phe Ile
    50                  55                  60

caa gca aat ttt aag aaa cga gaa tgt gtc ttc ttt acc aaa gat tcc      240
Gln Ala Asn Phe Lys Lys Arg Glu Cys Val Phe Phe Thr Lys Asp Ser
65                  70                  75                  80

aag gcc acg gag aat gtg tgc aag tgt ggc tat gcc cag agc cag cac      288
Lys Ala Thr Glu Asn Val Cys Lys Cys Gly Tyr Ala Gln Ser Gln His
                85                  90                  95

atg gaa ggc acc cag atc aac caa agt gag aaa tgg aac tac aag aaa      336
Met Glu Gly Thr Gln Ile Asn Gln Ser Glu Lys Trp Asn Tyr Lys Lys
            100                 105                 110

cac acc aag gaa ttt cct acc gac gcc ttt ggg gat att cag ttt gag      384
His Thr Lys Glu Phe Pro Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu
        115                 120                 125

aca ctg ggg aag aaa ggg aag tat ata cgt ctg tcc tgc gac acg gac      432
Thr Leu Gly Lys Lys Gly Lys Tyr Ile Arg Leu Ser Cys Asp Thr Asp
    130                 135                 140

gcg gaa atc ctt tac gag ctg ctg acc cag cac tgg cac ctg aaa aca      480
Ala Glu Ile Leu Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr
145                 150                 155                 160

ccc aac ctg gtc att tct gtg acc ggg ggc gcc aag aac ttc gcc ctg      528
Pro Asn Leu Val Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Ala Leu
                165                 170                 175

aag ccg cgc atg cgc aag atc ttc agc cgg ctc atc tac atc gcg cag      576
Lys Pro Arg Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln
            180                 185                 190

tcc aaa ggt gct tgg att ctc acg gga ggc acc cat tat ggc ctg atg      624
Ser Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His Tyr Gly Leu Met
        195                 200                 205

aag tac ctc ggg gag gtg gtg aga gat aac acc atc agc agg agt tca      672
Lys Tyr Leu Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser Ser
    210                 215                 220

gag gag aat att gtg gcc att ggc ata gca gct tgg ggc atg gtc tcc      720
Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Ala Trp Gly Met Val Ser
225                 230                 235                 240

aac cgg gac acc ctc atc agg aat tgc gat gct gag ggc tat ttt tta      768
Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly Tyr Phe Leu
                245                 250                 255
```

-continued

```
gcc cag tac ctt atg gat gac ttc aca aga gat cca ctg tat atc ctg      816
Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro Leu Tyr Ile Leu
            260                 265                 270

gac aac aac cac aca cat ttg ctg ctc gtg gac aat ggc tgt cat gga      864
Asp Asn Asn His Thr His Leu Leu Leu Val Asp Asn Gly Cys His Gly
        275                 280                 285

cat ccc act gtc gaa gca aag ctc cgg aat cag cta gag aag tat atc      912
His Pro Thr Val Glu Ala Lys Leu Arg Asn Gln Leu Glu Lys Tyr Ile
    290                 295                 300

tct gag cgc act att caa gat tcc aac tat ggt ggc aag atc ccc att      960
Ser Glu Arg Thr Ile Gln Asp Ser Asn Tyr Gly Gly Lys Ile Pro Ile
305                 310                 315                 320

gtg tgt ttt gcc caa gga ggt gga aaa gag act ttg aaa gcc atc aat     1008
Val Cys Phe Ala Gln Gly Gly Gly Lys Glu Thr Leu Lys Ala Ile Asn
                325                 330                 335

acc tcc atc aaa aat aaa att cct tgt gtg gtg gtg gaa ggc tcg ggc     1056
Thr Ser Ile Lys Asn Lys Ile Pro Cys Val Val Val Glu Gly Ser Gly
            340                 345                 350

cag atc gct gat gtg atc gct agc ctg gtg gag gtg gag gat gcc ctg     1104
Gln Ile Ala Asp Val Ile Ala Ser Leu Val Glu Val Glu Asp Ala Leu
        355                 360                 365

aca tct tct gcc gtc aag gag aag ctg gtg cgc ttt tta ccc cgc acg     1152
Thr Ser Ser Ala Val Lys Glu Lys Leu Val Arg Phe Leu Pro Arg Thr
    370                 375                 380

gtg tcc cgg ctg cct gag gag gag act gag agt tgg atc aaa tgg ctc     1200
Val Ser Arg Leu Pro Glu Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu
385                 390                 395                 400

aaa gaa att ctc gaa tgt tct cac cta tta aca gtt att aaa atg gaa     1248
Lys Glu Ile Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu
                405                 410                 415

gaa gct ggg gat gaa att gtg agc aat gcc atc tcc tac gct cta tac     1296
Glu Ala Gly Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr
            420                 425                 430

aaa gcc ttc agc acc agt gag caa gac aag gat aac tgg aat ggg cag     1344
Lys Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln
        435                 440                 445

ctg aag ctt ctg ctg gag tgg aac cag ctg gac tta gcc aat gat gag     1392
Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp Glu
    450                 455                 460

att ttc acc aat gac cgc cga tgg gag aag agc aaa ccg agg ctc aga     1440
Ile Phe Thr Asn Asp Arg Arg Trp Glu Lys Ser Lys Pro Arg Leu Arg
465                 470                 475                 480

gac aca ata atc cag gtc aca tgg ctg gaa aat ggt aga atc aag gtt     1488
Asp Thr Ile Ile Gln Val Thr Trp Leu Glu Asn Gly Arg Ile Lys Val
                485                 490                 495

gag agc aaa gat gtg act gac ggc aaa gcc tct tct cat atg ctg gtg     1536
Glu Ser Lys Asp Val Thr Asp Gly Lys Ala Ser Ser His Met Leu Val
            500                 505                 510

gtt ctc aag tct gct gac ctt caa gaa gtc atg ttt acg gct ctc ata     1584
Val Leu Lys Ser Ala Asp Leu Gln Glu Val Met Phe Thr Ala Leu Ile
        515                 520                 525

aag gac aga ccc aag ttt gtc cgc ctc ttt ctg gag aat ggc ttg aac     1632
Lys Asp Arg Pro Lys Phe Val Arg Leu Phe Leu Glu Asn Gly Leu Asn
    530                 535                 540

cta cgg aag ttt ctc acc cat gat gtc ctc act gaa ctc ttc tcc aac     1680
Leu Arg Lys Phe Leu Thr His Asp Val Leu Thr Glu Leu Phe Ser Asn
545                 550                 555                 560
```

-continued

```
cac ttc agc acg ctt gtg tac cgg aat ctg cag atc gcc aag aat tcc     1728
His Phe Ser Thr Leu Val Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser
                565                 570                 575

tat aat gat gcc ctc ctc acg ttt gtc tgg aaa ctg gtt gcg aac ttc     1776
Tyr Asn Asp Ala Leu Leu Thr Phe Val Trp Lys Leu Val Ala Asn Phe
            580                 585                 590

cga aga ggc ttc cgg aag gaa gac aga aat ggc cgg gac gag atg gac     1824
Arg Arg Gly Phe Arg Lys Glu Asp Arg Asn Gly Arg Asp Glu Met Asp
        595                 600                 605

ata gaa ctc cac gac gtg tct cct att act cgg cac ccc ctg caa gct     1872
Ile Glu Leu His Asp Val Ser Pro Ile Thr Arg His Pro Leu Gln Ala
    610                 615                 620

ctc ttc atc tgg gcc att ctt cag aat aag aag gaa ctc tcc aaa gtc     1920
Leu Phe Ile Trp Ala Ile Leu Gln Asn Lys Lys Glu Leu Ser Lys Val
625                 630                 635                 640

att tgg gag cag acc agg ggc tgc act ctg gca gcc ctg gga gcc agc     1968
Ile Trp Glu Gln Thr Arg Gly Cys Thr Leu Ala Ala Leu Gly Ala Ser
                645                 650                 655

aag ctt ctg aag act ctg gcc aaa gtg aag aac gac atc aat gct gct     2016
Lys Leu Leu Lys Thr Leu Ala Lys Val Lys Asn Asp Ile Asn Ala Ala
            660                 665                 670

ggg gag tcc gag gag ctg gct aat gag tac ctg acc cgg gct gtt ggt     2064
Gly Glu Ser Glu Glu Leu Ala Asn Glu Tyr Leu Thr Arg Ala Val Gly
        675                 680                 685

gag tcc aca gtg tgg aat gct gtg gtg ggc gcg gat ctg cca tgt ggc     2112
Glu Ser Thr Val Trp Asn Ala Val Val Gly Ala Asp Leu Pro Cys Gly
    690                 695                 700

aca gac att gcc agc ggc act cat aga cca gat ggt gga gag ctg ttc     2160
Thr Asp Ile Ala Ser Gly Thr His Arg Pro Asp Gly Gly Glu Leu Phe
705                 710                 715                 720

act gag tgt tac agc agc gat gaa gac ttg gca gaa cag ctg ctg gtc     2208
Thr Glu Cys Tyr Ser Ser Asp Glu Asp Leu Ala Glu Gln Leu Leu Val
                725                 730                 735

tat tcc tgt gaa gct tgg ggt gga agc aac tgt ctg gag ctg gcg gtg     2256
Tyr Ser Cys Glu Ala Trp Gly Gly Ser Asn Cys Leu Glu Leu Ala Val
            740                 745                 750

gag gcc aca gac cag cat ttc atc gcc cag cct ggg gtc cag aat ttt     2304
Glu Ala Thr Asp Gln His Phe Ile Ala Gln Pro Gly Val Gln Asn Phe
        755                 760                 765

ctt tct aag caa tgg tat gga gag att tcc cga gac acc aag aac tgg     2352
Leu Ser Lys Gln Trp Tyr Gly Glu Ile Ser Arg Asp Thr Lys Asn Trp
    770                 775                 780

aag att atc ctg tgt ctg ttt att ata ccc ttg gtg ggc tgt ggc ttt     2400
Lys Ile Ile Leu Cys Leu Phe Ile Ile Pro Leu Val Gly Cys Gly Phe
785                 790                 795                 800

gta tca ttt agg aag aaa cct gtc gac aag cac aag aag ctg ctt tgg     2448
Val Ser Phe Arg Lys Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp
                805                 810                 815

tac tat gtg gcg ttc ttc acc tcc ccc ttc gtg gtc ttc tcc tgg aat     2496
Tyr Tyr Val Ala Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn
            820                 825                 830

gtg gtc ttc tac atc gcc ttc ctc ctg ctg ttt gcc tac gtg ctg ctc     2544
Val Val Phe Tyr Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu
        835                 840                 845

atg gat ttc cat tcg gtg cca cac ccc ccc gag ctg gtc ctg tac tcg     2592
Met Asp Phe His Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser
    850                 855                 860
```

-continued

```
ctg gtc ttt gtc ctc ttc tgt gat gaa gtg aga cag ggc cgg ccg gct     2640
Leu Val Phe Val Leu Phe Cys Asp Glu Val Arg Gln Gly Arg Pro Ala
865                 870                 875                 880

gct ccc agt gcg ggg ccc gcc aag ccc acg ccc acc cgg aac tcc atc     2688
Ala Pro Ser Ala Gly Pro Ala Lys Pro Thr Pro Thr Arg Asn Ser Ile
                885                 890                 895

tgg ccc gca agc tcc aca cgc agc ccc ggt tcc cgc tca cgc cac tcc     2736
Trp Pro Ala Ser Ser Thr Arg Ser Pro Gly Ser Arg Ser Arg His Ser
            900                 905                 910

ttc cac act tcc ctg caa gct gag ggt gcc agc tct ggc ctt ggc cag     2784
Phe His Thr Ser Leu Gln Ala Glu Gly Ala Ser Ser Gly Leu Gly Gln
        915                 920                 925

ccc aga aag ggg ctc cca cag tgc agc ggt ggg ctg aag ggc tcc tca     2832
Pro Arg Lys Gly Leu Pro Gln Cys Ser Gly Gly Leu Lys Gly Ser Ser
    930                 935                 940

agt gcc gcc aaa gtg gga gcc cag gca gag gag gtg ccg aga gca agc     2880
Ser Ala Ala Lys Val Gly Ala Gln Ala Glu Glu Val Pro Arg Ala Ser
945                 950                 955                 960

gag ggc tgt gag gac tgc cag cac gct gtc acc tct cag aag cgt aag     2928
Glu Gly Cys Glu Asp Cys Gln His Ala Val Thr Ser Gln Lys Arg Lys
                965                 970                 975

aca gca atg gac caa aca gac gaa gat ctc ttc ccc tat gga gca ttc     2976
Thr Ala Met Asp Gln Thr Asp Glu Asp Leu Phe Pro Tyr Gly Ala Phe
            980                 985                 990

tac cag ttc ctg atg att tcc agg agc ttt cga gga gag gag atg agc     3024
Tyr Gln Phe Leu Met Ile Ser Arg Ser Phe Arg Gly Glu Glu Met Ser
        995                 1000                1005

atc ggc aag cag cac taa                                             3042
Ile Gly Lys Gln His  *
    1010


<210> SEQ ID NO 27
<211> LENGTH: 3448
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (425)...(2971)

<400> SEQUENCE: 27

attaaccttc tcttagtctt caacctaagt acttgaatgt caagtaccct ccaaccctca     60

atgtcccaag acttttaaga gcggaaggta ccgatgagtt ccatccttta ctagggtcac    120

caaggaaggc atgggtatat ggaaattttt attattattc catctgaata tcattttcta    180

gagaatagga gcttttgttc tgaagggctg ccggcttcct tctgggatct agcagccagg    240

gttagatcac aggtgtcact ttcaggcgag tagttagcaa cggtatcgct agcaactgag    300

ccgacccctg cagccagagg tttgcagtgg gtagtgtgta ttccagaaag ggccctgaca    360

tgtgaaagga aggaatgtgc cctaatattc tacagttgtt ttatcgttgc tactgattag    420

gtcc atg gag gga agc cca tcc ctg aga cgc atg aca gtg atg cgg gag     469
Met Glu Gly Ser Pro Ser Leu Arg Arg Met Thr Val Met Arg Glu
 1               5                  10                  15

aag ggc cgg cgc cag gct gtc agg ggc ccg gcc ttc atg ttc aat gac      517
Lys Gly Arg Arg Gln Ala Val Arg Gly Pro Ala Phe Met Phe Asn Asp
                20                  25                  30

cgc ggc acc agc ctc acc gcc gag gag gag cgc ttc ctc gac gcc gcc      565
Arg Gly Thr Ser Leu Thr Ala Glu Glu Glu Arg Phe Leu Asp Ala Ala
```

-continued

```
            35                  40                  45

gag tac ggc aac atc cca gtg gtg cgc aag atg ctg gag gag tcc aag      613
Glu Tyr Gly Asn Ile Pro Val Val Arg Lys Met Leu Glu Glu Ser Lys
        50                  55                  60

acg ctg aac gtc aac tgc gtg gac tac atg ggc cag aac gcg ctg cag      661
Thr Leu Asn Val Asn Cys Val Asp Tyr Met Gly Gln Asn Ala Leu Gln
    65                  70                  75

ctg gct gtg ggc aac gag cac ctg gag gtg acc gag ctg ctg ctc aag      709
Leu Ala Val Gly Asn Glu His Leu Glu Val Thr Glu Leu Leu Leu Lys
80                  85                  90                  95

aag gag aac ctg gcg cgc att ggc gac gcc ctg ctg ctc gcc atc agc      757
Lys Glu Asn Leu Ala Arg Ile Gly Asp Ala Leu Leu Leu Ala Ile Ser
                100                 105                 110

aag ggc tac gtg cgc atc gta gag gcc atc ctc aac cac cct ggc ttc      805
Lys Gly Tyr Val Arg Ile Val Glu Ala Ile Leu Asn His Pro Gly Phe
            115                 120                 125

gcg gcc agc aag cgt ctc act ctg agc ccc tgt gag cag gag ctg cag      853
Ala Ala Ser Lys Arg Leu Thr Leu Ser Pro Cys Glu Gln Glu Leu Gln
        130                 135                 140

gac gac gac ttc tac gct tac gac gag gac ggc acg cgc ttc tcg ccg      901
Asp Asp Asp Phe Tyr Ala Tyr Asp Glu Asp Gly Thr Arg Phe Ser Pro
    145                 150                 155

gac atc acc ccc atc atc ctg gcg gcg cac tgc cag aaa tac gaa gtg      949
Asp Ile Thr Pro Ile Ile Leu Ala Ala His Cys Gln Lys Tyr Glu Val
160                 165                 170                 175

gtg cac atg ctg ctg atg aag ggt gcc agg atc gag cgg ccg cac gac      997
Val His Met Leu Leu Met Lys Gly Ala Arg Ile Glu Arg Pro His Asp
                180                 185                 190

tat ttc tgc aag tgc ggg gac tgc atg gag aag cag agg cac gac tcc     1045
Tyr Phe Cys Lys Cys Gly Asp Cys Met Glu Lys Gln Arg His Asp Ser
            195                 200                 205

ttc agc cac tca cgc tcg agg atc aat gcc tac aag ggg ctg gcc agc     1093
Phe Ser His Ser Arg Ser Arg Ile Asn Ala Tyr Lys Gly Leu Ala Ser
        210                 215                 220

ccg gct tac ctc tca ttg tcc agc gag gac ccg gtg ctt acg gcc cta     1141
Pro Ala Tyr Leu Ser Leu Ser Ser Glu Asp Pro Val Leu Thr Ala Leu
    225                 230                 235

gag ctc agc aac gag ctg gcc aag ctg gcc aac ata gag aag gag ttc     1189
Glu Leu Ser Asn Glu Leu Ala Lys Leu Ala Asn Ile Glu Lys Glu Phe
240                 245                 250                 255

aag aat gac tat cgg aag ctc tcc atg caa tgc aaa gac ttt gta gtg     1237
Lys Asn Asp Tyr Arg Lys Leu Ser Met Gln Cys Lys Asp Phe Val Val
                260                 265                 270

ggt gtg ctg gat ctc tgc cga gac tca gaa gag gta gaa gcc att ctg     1285
Gly Val Leu Asp Leu Cys Arg Asp Ser Glu Glu Val Glu Ala Ile Leu
            275                 280                 285

aat gga gat ctg gaa tca gca gag cct ctg gag gta cac agg cac aaa     1333
Asn Gly Asp Leu Glu Ser Ala Glu Pro Leu Glu Val His Arg His Lys
        290                 295                 300

gct tca tta agt cgt gtc aaa ctt gcc att aag tat gaa gtc aaa aag     1381
Ala Ser Leu Ser Arg Val Lys Leu Ala Ile Lys Tyr Glu Val Lys Lys
    305                 310                 315

ttt gtg gct cat ccc aac tgc cag cag cag ctc ttg acg atc tgg tat     1429
Phe Val Ala His Pro Asn Cys Gln Gln Gln Leu Leu Thr Ile Trp Tyr
320                 325                 330                 335

gag aac ctc tca ggc cta agg gag cag acc ata gct atc aag tgt ctc     1477
Glu Asn Leu Ser Gly Leu Arg Glu Gln Thr Ile Ala Ile Lys Cys Leu
```

-continued

```
                340                 345                 350

gtt gtg ctg gtc gtg gcc ctg ggc ctt cca ttc ctg gcc att ggc tac     1525
Val Val Leu Val Val Ala Leu Gly Leu Pro Phe Leu Ala Ile Gly Tyr
            355                 360                 365

tgg atc gca cct tgc agc agg ctg ggg aaa att ctg cga agc cct ttt     1573
Trp Ile Ala Pro Cys Ser Arg Leu Gly Lys Ile Leu Arg Ser Pro Phe
        370                 375                 380

atg aag ttt gta gca cat gca gct tct ttc atc atc ttc ctg ggt ctg     1621
Met Lys Phe Val Ala His Ala Ala Ser Phe Ile Ile Phe Leu Gly Leu
    385                 390                 395

ctt gtg ttc aat gcc tca gac agg ttc gaa ggc atc acc acg ctg ccc     1669
Leu Val Phe Asn Ala Ser Asp Arg Phe Glu Gly Ile Thr Thr Leu Pro
400                 405                 410                 415

aat atc aca gtt act gac tat ccc aaa cag atc ttc agg gtg aaa acc     1717
Asn Ile Thr Val Thr Asp Tyr Pro Lys Gln Ile Phe Arg Val Lys Thr
                420                 425                 430

acc cag ttt aca tgg act gaa atg cta att atg gtc tgg gtt ctt gga     1765
Thr Gln Phe Thr Trp Thr Glu Met Leu Ile Met Val Trp Val Leu Gly
            435                 440                 445

atg atg tgg tct gaa tgt aaa gag ctc tgg ctg gaa gga cct agg gaa     1813
Met Met Trp Ser Glu Cys Lys Glu Leu Trp Leu Glu Gly Pro Arg Glu
        450                 455                 460

tac att ttg cag ttg tgg aat gtg ctt gac ttt ggg atg ctg tcc atc     1861
Tyr Ile Leu Gln Leu Trp Asn Val Leu Asp Phe Gly Met Leu Ser Ile
    465                 470                 475

ttc att gct gct ttc aca gcc aga ttc cta gct ttc ctt cag gca acg     1909
Phe Ile Ala Ala Phe Thr Ala Arg Phe Leu Ala Phe Leu Gln Ala Thr
480                 485                 490                 495

aag gca caa cag tat gtg gac agt tac gtc caa gag agt gac ctc agt     1957
Lys Ala Gln Gln Tyr Val Asp Ser Tyr Val Gln Glu Ser Asp Leu Ser
                500                 505                 510

gaa gtg aca ctc cca cca gag ata cag tat ttc act tat gct aga gat     2005
Glu Val Thr Leu Pro Pro Glu Ile Gln Tyr Phe Thr Tyr Ala Arg Asp
            515                 520                 525

aaa tgg ctc cct tct gac cct cag att ata tct gaa ggc ctt tat gcc     2053
Lys Trp Leu Pro Ser Asp Pro Gln Ile Ile Ser Glu Gly Leu Tyr Ala
        530                 535                 540

ata gct gtt gtg ctc agc ttc tct cgg att gcg tac atc ctc cct gca     2101
Ile Ala Val Val Leu Ser Phe Ser Arg Ile Ala Tyr Ile Leu Pro Ala
    545                 550                 555

aat gag agc ttt ggc ccc ctg cag atc tct ctt gga agg act gta aag     2149
Asn Glu Ser Phe Gly Pro Leu Gln Ile Ser Leu Gly Arg Thr Val Lys
560                 565                 570                 575

gac ata ttc aag ttc atg gtc ctc ttt att atg gtg ttt ttt gcc ttt     2197
Asp Ile Phe Lys Phe Met Val Leu Phe Ile Met Val Phe Phe Ala Phe
                580                 585                 590

atg att ggc atg ttc ata ctt tat tct tac tac ctt ggg gct aaa gtt     2245
Met Ile Gly Met Phe Ile Leu Tyr Ser Tyr Tyr Leu Gly Ala Lys Val
            595                 600                 605

aat gct gct ttt acc act gta gaa gaa agt ttc aag act tta ttt tgg     2293
Asn Ala Ala Phe Thr Thr Val Glu Glu Ser Phe Lys Thr Leu Phe Trp
        610                 615                 620

tca ata ttt ggg ttg tct gaa gtg act tcc gtt gtg ctc aaa tat gat     2341
Ser Ile Phe Gly Leu Ser Glu Val Thr Ser Val Val Leu Lys Tyr Asp
    625                 630                 635

cac aaa ttc ata gaa aat att gga tac gtt ctt tat gga ata tac aat     2389
His Lys Phe Ile Glu Asn Ile Gly Tyr Val Leu Tyr Gly Ile Tyr Asn
```

-continued

```
640                 645                 650                 655

gta act atg gtg gtc gtt tta ctc aac atg cta att gct atg att aat     2437
Val Thr Met Val Val Val Leu Leu Asn Met Leu Ile Ala Met Ile Asn
                660                 665                 670

agc tca tat caa gaa att gag gat gac agt gat gta gaa tgg aag ttt     2485
Ser Ser Tyr Gln Glu Ile Glu Asp Asp Ser Asp Val Glu Trp Lys Phe
            675                 680                 685

gct cgt tca aaa ctt tgg tta tcc tat ttt gat gat gga aaa aca tta     2533
Ala Arg Ser Lys Leu Trp Leu Ser Tyr Phe Asp Asp Gly Lys Thr Leu
        690                 695                 700

cct cca cct ttc agt cta gtt cct agt cca aaa tca ttt gtt tat ttc     2581
Pro Pro Pro Phe Ser Leu Val Pro Ser Pro Lys Ser Phe Val Tyr Phe
    705                 710                 715

atc atg cga att gtt aac ttt ccc aaa tgc aga agg aga agg ctt cag     2629
Ile Met Arg Ile Val Asn Phe Pro Lys Cys Arg Arg Arg Arg Leu Gln
720                 725                 730                 735

aag gat ata gaa atg gga atg ggt aac tca aag tcc agg tta aac ctc     2677
Lys Asp Ile Glu Met Gly Met Gly Asn Ser Lys Ser Arg Leu Asn Leu
                740                 745                 750

ttc act cag tct aac tca aga gtt ttt gaa tca cac agt ttt aac agc     2725
Phe Thr Gln Ser Asn Ser Arg Val Phe Glu Ser His Ser Phe Asn Ser
            755                 760                 765

att ctc aat cag cca aca cgt tat cag cag ata atg aaa aga ctt ata     2773
Ile Leu Asn Gln Pro Thr Arg Tyr Gln Gln Ile Met Lys Arg Leu Ile
        770                 775                 780

aag cgg tat gtt ttg aaa gca caa gta gac aaa gaa aat gat gaa gtt     2821
Lys Arg Tyr Val Leu Lys Ala Gln Val Asp Lys Glu Asn Asp Glu Val
    785                 790                 795

aat gaa ggt gaa tta aaa gaa atc aag caa gat atc tcc agc ctt cgt     2869
Asn Glu Gly Glu Leu Lys Glu Ile Lys Gln Asp Ile Ser Ser Leu Arg
800                 805                 810                 815

tat gaa ctt ttg gaa gac aag agc caa gca act gag gaa tta gcc att     2917
Tyr Glu Leu Leu Glu Asp Lys Ser Gln Ala Thr Glu Glu Leu Ala Ile
                820                 825                 830

cta att cat aaa ctt agt gag aaa ctg aat ccc agc atg ctg aga tgt     2965
Leu Ile His Lys Leu Ser Glu Lys Leu Asn Pro Ser Met Leu Arg Cys
            835                 840                 845

gaa tga tgcagcaacc tggatttggc tttgactata gcacaaatgt gggcaataat      3021
Glu  *

atttctaagt atgaaatact tgaaaaacta tgatgtaaat ttttagtatt aactaccttt   3081

atcatgtgaa cctttaaaag ttagctctta atggttttat tgttttatca catgaaaatg   3141

cattttattt gtctgctttg acattacagt ggcataccat tgtgttgaaa agcccaatat   3201

tactatatta ttgaaacttt tattcatttt agagtaaact ccacatcttt gcactacctg   3261

tttgcctcca agagactatc agttccttgg ggacagggac catgtcttat tcatctttgt   3321

gtctccagca tctagtacag tgcctggtat atagtaggtg ctcaataaat gttgaaacca   3381

actgaactgc caacaaaata aaaataaaaa gtcttcacta tgtagcataa aaaaaaaaaa   3441

aaaaaaa                                                             3448
```

<210> SEQ ID NO 28
<211> LENGTH: 848
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 28

-continued

```
Met Glu Gly Ser Pro Ser Leu Arg Arg Met Thr Val Met Arg Glu Lys
 1               5                  10                  15

Gly Arg Arg Gln Ala Val Arg Gly Pro Ala Phe Met Phe Asn Asp Arg
            20                  25                  30

Gly Thr Ser Leu Thr Ala Glu Glu Glu Arg Phe Leu Asp Ala Ala Glu
        35                  40                  45

Tyr Gly Asn Ile Pro Val Val Arg Lys Met Leu Glu Glu Ser Lys Thr
    50                  55                  60

Leu Asn Val Asn Cys Val Asp Tyr Met Gly Gln Asn Ala Leu Gln Leu
65                  70                  75                  80

Ala Val Gly Asn Glu His Leu Glu Val Thr Glu Leu Leu Leu Lys Lys
                85                  90                  95

Glu Asn Leu Ala Arg Ile Gly Asp Ala Leu Leu Leu Ala Ile Ser Lys
            100                 105                 110

Gly Tyr Val Arg Ile Val Glu Ala Ile Leu Asn His Pro Gly Phe Ala
        115                 120                 125

Ala Ser Lys Arg Leu Thr Leu Ser Pro Cys Glu Gln Glu Leu Gln Asp
    130                 135                 140

Asp Asp Phe Tyr Ala Tyr Asp Glu Asp Gly Thr Arg Phe Ser Pro Asp
145                 150                 155                 160

Ile Thr Pro Ile Ile Leu Ala Ala His Cys Gln Lys Tyr Glu Val Val
                165                 170                 175

His Met Leu Leu Met Lys Gly Ala Arg Ile Glu Arg Pro His Asp Tyr
            180                 185                 190

Phe Cys Lys Cys Gly Asp Cys Met Glu Lys Gln Arg His Asp Ser Phe
        195                 200                 205

Ser His Ser Arg Ser Arg Ile Asn Ala Tyr Lys Gly Leu Ala Ser Pro
    210                 215                 220

Ala Tyr Leu Ser Leu Ser Ser Glu Asp Pro Val Leu Thr Ala Leu Glu
225                 230                 235                 240

Leu Ser Asn Glu Leu Ala Lys Leu Ala Asn Ile Glu Lys Glu Phe Lys
                245                 250                 255

Asn Asp Tyr Arg Lys Leu Ser Met Gln Cys Lys Asp Phe Val Val Gly
            260                 265                 270

Val Leu Asp Leu Cys Arg Asp Ser Glu Glu Val Glu Ala Ile Leu Asn
        275                 280                 285

Gly Asp Leu Glu Ser Ala Glu Pro Leu Glu Val His Arg His Lys Ala
    290                 295                 300

Ser Leu Ser Arg Val Lys Leu Ala Ile Lys Tyr Glu Val Lys Lys Phe
305                 310                 315                 320

Val Ala His Pro Asn Cys Gln Gln Gln Leu Leu Thr Ile Trp Tyr Glu
                325                 330                 335

Asn Leu Ser Gly Leu Arg Glu Gln Thr Ile Ala Ile Lys Cys Leu Val
            340                 345                 350

Val Leu Val Val Ala Leu Gly Leu Pro Phe Leu Ala Ile Gly Tyr Trp
        355                 360                 365

Ile Ala Pro Cys Ser Arg Leu Gly Lys Ile Leu Arg Ser Pro Phe Met
    370                 375                 380

Lys Phe Val Ala His Ala Ala Ser Phe Ile Ile Phe Leu Gly Leu Leu
385                 390                 395                 400
```

-continued

```
Val Phe Asn Ala Ser Asp Arg Phe Glu Gly Ile Thr Thr Leu Pro Asn
                405                 410                 415

Ile Thr Val Thr Asp Tyr Pro Lys Gln Ile Phe Arg Val Lys Thr Thr
            420                 425                 430

Gln Phe Thr Trp Thr Glu Met Leu Ile Met Val Trp Val Leu Gly Met
        435                 440                 445

Met Trp Ser Glu Cys Lys Glu Leu Trp Leu Glu Gly Pro Arg Glu Tyr
    450                 455                 460

Ile Leu Gln Leu Trp Asn Val Leu Asp Phe Gly Met Leu Ser Ile Phe
465                 470                 475                 480

Ile Ala Ala Phe Thr Ala Arg Phe Leu Ala Phe Leu Gln Ala Thr Lys
                485                 490                 495

Ala Gln Gln Tyr Val Asp Ser Tyr Val Gln Glu Ser Asp Leu Ser Glu
            500                 505                 510

Val Thr Leu Pro Pro Glu Ile Gln Tyr Phe Thr Tyr Ala Arg Asp Lys
        515                 520                 525

Trp Leu Pro Ser Asp Pro Gln Ile Ile Ser Glu Gly Leu Tyr Ala Ile
    530                 535                 540

Ala Val Val Leu Ser Phe Ser Arg Ile Ala Tyr Ile Leu Pro Ala Asn
545                 550                 555                 560

Glu Ser Phe Gly Pro Leu Gln Ile Ser Leu Gly Arg Thr Val Lys Asp
                565                 570                 575

Ile Phe Lys Phe Met Val Leu Phe Ile Met Val Phe Phe Ala Phe Met
            580                 585                 590

Ile Gly Met Phe Ile Leu Tyr Ser Tyr Tyr Leu Gly Ala Lys Val Asn
        595                 600                 605

Ala Ala Phe Thr Thr Val Glu Glu Ser Phe Lys Thr Leu Phe Trp Ser
    610                 615                 620

Ile Phe Gly Leu Ser Glu Val Thr Ser Val Val Leu Lys Tyr Asp His
625                 630                 635                 640

Lys Phe Ile Glu Asn Ile Gly Tyr Val Leu Tyr Gly Ile Tyr Asn Val
                645                 650                 655

Thr Met Val Val Val Leu Leu Asn Met Leu Ile Ala Met Ile Asn Ser
            660                 665                 670

Ser Tyr Gln Glu Ile Glu Asp Asp Ser Asp Val Glu Trp Lys Phe Ala
        675                 680                 685

Arg Ser Lys Leu Trp Leu Ser Tyr Phe Asp Asp Gly Lys Thr Leu Pro
    690                 695                 700

Pro Pro Phe Ser Leu Val Pro Ser Pro Lys Ser Phe Val Tyr Phe Ile
705                 710                 715                 720

Met Arg Ile Val Asn Phe Pro Lys Cys Arg Arg Arg Arg Leu Gln Lys
                725                 730                 735

Asp Ile Glu Met Gly Met Gly Asn Ser Lys Ser Arg Leu Asn Leu Phe
            740                 745                 750

Thr Gln Ser Asn Ser Arg Val Phe Glu Ser His Ser Phe Asn Ser Ile
        755                 760                 765

Leu Asn Gln Pro Thr Arg Tyr Gln Gln Ile Met Lys Arg Leu Ile Lys
    770                 775                 780

Arg Tyr Val Leu Lys Ala Gln Val Asp Lys Glu Asn Asp Glu Val Asn
785                 790                 795                 800

Glu Gly Glu Leu Lys Glu Ile Lys Gln Asp Ile Ser Ser Leu Arg Tyr
```

-continued

```
                805                 810                 815

Glu Leu Leu Glu Asp Lys Ser Gln Ala Thr Glu Glu Leu Ala Ile Leu
            820                 825                 830

Ile His Lys Leu Ser Glu Lys Leu Asn Pro Ser Met Leu Arg Cys Glu
        835                 840                 845


<210> SEQ ID NO 29
<211> LENGTH: 2547
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(2547)

<400> SEQUENCE: 29

atg gag gga agc cca tcc ctg aga cgc atg aca gtg atg cgg gag aag       48
Met Glu Gly Ser Pro Ser Leu Arg Arg Met Thr Val Met Arg Glu Lys
 1               5                  10                  15

ggc cgg cgc cag gct gtc agg ggc ccg gcc ttc atg ttc aat gac cgc       96
Gly Arg Arg Gln Ala Val Arg Gly Pro Ala Phe Met Phe Asn Asp Arg
            20                  25                  30

ggc acc agc ctc acc gcc gag gag gag cgc ttc ctc gac gcc gcc gag      144
Gly Thr Ser Leu Thr Ala Glu Glu Glu Arg Phe Leu Asp Ala Ala Glu
        35                  40                  45

tac ggc aac atc cca gtg gtg cgc aag atg ctg gag gag tcc aag acg      192
Tyr Gly Asn Ile Pro Val Val Arg Lys Met Leu Glu Glu Ser Lys Thr
    50                  55                  60

ctg aac gtc aac tgc gtg gac tac atg ggc cag aac gcg ctg cag ctg      240
Leu Asn Val Asn Cys Val Asp Tyr Met Gly Gln Asn Ala Leu Gln Leu
 65                  70                  75                  80

gct gtg ggc aac gag cac ctg gag gtg acc gag ctg ctg ctc aag aag      288
Ala Val Gly Asn Glu His Leu Glu Val Thr Glu Leu Leu Leu Lys Lys
                85                  90                  95

gag aac ctg gcg cgc att ggc gac gcc ctg ctg ctc gcc atc agc aag      336
Glu Asn Leu Ala Arg Ile Gly Asp Ala Leu Leu Leu Ala Ile Ser Lys
            100                 105                 110

ggc tac gtg cgc atc gta gag gcc atc ctc aac cac cct ggc ttc gcg      384
Gly Tyr Val Arg Ile Val Glu Ala Ile Leu Asn His Pro Gly Phe Ala
        115                 120                 125

gcc agc aag cgt ctc act ctg agc ccc tgt gag cag gag ctg cag gac      432
Ala Ser Lys Arg Leu Thr Leu Ser Pro Cys Glu Gln Glu Leu Gln Asp
    130                 135                 140

gac gac ttc tac gct tac gac gag gac ggc acg cgc ttc tcg ccg gac      480
Asp Asp Phe Tyr Ala Tyr Asp Glu Asp Gly Thr Arg Phe Ser Pro Asp
145                 150                 155                 160

atc acc ccc atc atc ctg gcg gcg cac tgc cag aaa tac gaa gtg gtg      528
Ile Thr Pro Ile Ile Leu Ala Ala His Cys Gln Lys Tyr Glu Val Val
                165                 170                 175

cac atg ctg ctg atg aag ggt gcc agg atc gag cgg ccg cac gac tat      576
His Met Leu Leu Met Lys Gly Ala Arg Ile Glu Arg Pro His Asp Tyr
            180                 185                 190

ttc tgc aag tgc ggg gac tgc atg gag aag cag agg cac gac tcc ttc      624
Phe Cys Lys Cys Gly Asp Cys Met Glu Lys Gln Arg His Asp Ser Phe
        195                 200                 205

agc cac tca cgc tcg agg atc aat gcc tac aag ggg ctg gcc agc ccg      672
Ser His Ser Arg Ser Arg Ile Asn Ala Tyr Lys Gly Leu Ala Ser Pro
    210                 215                 220

gct tac ctc tca ttg tcc agc gag gac ccg gtg ctt acg gcc cta gag      720
```

-continued

```
Ala Tyr Leu Ser Leu Ser Ser Glu Asp Pro Val Leu Thr Ala Leu Glu
225                 230                 235                 240

ctc agc aac gag ctg gcc aag ctg gcc aac ata gag aag gag ttc aag      768
Leu Ser Asn Glu Leu Ala Lys Leu Ala Asn Ile Glu Lys Glu Phe Lys
                245                 250                 255

aat gac tat cgg aag ctc tcc atg caa tgc aaa gac ttt gta gtg ggt      816
Asn Asp Tyr Arg Lys Leu Ser Met Gln Cys Lys Asp Phe Val Val Gly
            260                 265                 270

gtg ctg gat ctc tgc cga gac tca gaa gag gta gaa gcc att ctg aat      864
Val Leu Asp Leu Cys Arg Asp Ser Glu Glu Val Glu Ala Ile Leu Asn
        275                 280                 285

gga gat ctg gaa tca gca gag cct ctg gag gta cac agg cac aaa gct      912
Gly Asp Leu Glu Ser Ala Glu Pro Leu Glu Val His Arg His Lys Ala
    290                 295                 300

tca tta agt cgt gtc aaa ctt gcc att aag tat gaa gtc aaa aag ttt      960
Ser Leu Ser Arg Val Lys Leu Ala Ile Lys Tyr Glu Val Lys Lys Phe
305                 310                 315                 320

gtg gct cat ccc aac tgc cag cag cag ctc ttg acg atc tgg tat gag     1008
Val Ala His Pro Asn Cys Gln Gln Gln Leu Leu Thr Ile Trp Tyr Glu
                325                 330                 335

aac ctc tca ggc cta agg gag cag acc ata gct atc aag tgt ctc gtt     1056
Asn Leu Ser Gly Leu Arg Glu Gln Thr Ile Ala Ile Lys Cys Leu Val
            340                 345                 350

gtg ctg gtc gtg gcc ctg ggc ctt cca ttc ctg gcc att ggc tac tgg     1104
Val Leu Val Val Ala Leu Gly Leu Pro Phe Leu Ala Ile Gly Tyr Trp
        355                 360                 365

atc gca cct tgc agc agg ctg ggg aaa att ctg cga agc cct ttt atg     1152
Ile Ala Pro Cys Ser Arg Leu Gly Lys Ile Leu Arg Ser Pro Phe Met
    370                 375                 380

aag ttt gta gca cat gca gct tct ttc atc atc ttc ctg ggt ctg ctt     1200
Lys Phe Val Ala His Ala Ala Ser Phe Ile Ile Phe Leu Gly Leu Leu
385                 390                 395                 400

gtg ttc aat gcc tca gac agg ttc gaa ggc atc acc acg ctg ccc aat     1248
Val Phe Asn Ala Ser Asp Arg Phe Glu Gly Ile Thr Thr Leu Pro Asn
                405                 410                 415

atc aca gtt act gac tat ccc aaa cag atc ttc agg gtg aaa acc acc     1296
Ile Thr Val Thr Asp Tyr Pro Lys Gln Ile Phe Arg Val Lys Thr Thr
            420                 425                 430

cag ttt aca tgg act gaa atg cta att atg gtc tgg gtt ctt gga atg     1344
Gln Phe Thr Trp Thr Glu Met Leu Ile Met Val Trp Val Leu Gly Met
        435                 440                 445

atg tgg tct gaa tgt aaa gag ctc tgg ctg gaa gga cct agg gaa tac     1392
Met Trp Ser Glu Cys Lys Glu Leu Trp Leu Glu Gly Pro Arg Glu Tyr
    450                 455                 460

att ttg cag ttg tgg aat gtg ctt gac ttt ggg atg ctg tcc atc ttc     1440
Ile Leu Gln Leu Trp Asn Val Leu Asp Phe Gly Met Leu Ser Ile Phe
465                 470                 475                 480

att gct gct ttc aca gcc aga ttc cta gct ttc ctt cag gca acg aag     1488
Ile Ala Ala Phe Thr Ala Arg Phe Leu Ala Phe Leu Gln Ala Thr Lys
                485                 490                 495

gca caa cag tat gtg gac agt tac gtc caa gag agt gac ctc agt gaa     1536
Ala Gln Gln Tyr Val Asp Ser Tyr Val Gln Glu Ser Asp Leu Ser Glu
            500                 505                 510

gtg aca ctc cca cca gag ata cag tat ttc act tat gct aga gat aaa     1584
Val Thr Leu Pro Pro Glu Ile Gln Tyr Phe Thr Tyr Ala Arg Asp Lys
        515                 520                 525

tgg ctc cct tct gac cct cag att ata tct gaa ggc ctt tat gcc ata     1632
```

-continued

```
Trp Leu Pro Ser Asp Pro Gln Ile Ile Ser Glu Gly Leu Tyr Ala Ile
    530                 535                 540

gct gtt gtg ctc agc ttc tct cgg att gcg tac atc ctc cct gca aat     1680
Ala Val Val Leu Ser Phe Ser Arg Ile Ala Tyr Ile Leu Pro Ala Asn
545                 550                 555                 560

gag agc ttt ggc ccc ctg cag atc tct ctt gga agg act gta aag gac     1728
Glu Ser Phe Gly Pro Leu Gln Ile Ser Leu Gly Arg Thr Val Lys Asp
                565                 570                 575

ata ttc aag ttc atg gtc ctc ttt att atg gtg ttt ttt gcc ttt atg     1776
Ile Phe Lys Phe Met Val Leu Phe Ile Met Val Phe Phe Ala Phe Met
            580                 585                 590

att ggc atg ttc ata ctt tat tct tac tac ctt ggg gct aaa gtt aat     1824
Ile Gly Met Phe Ile Leu Tyr Ser Tyr Tyr Leu Gly Ala Lys Val Asn
        595                 600                 605

gct gct ttt acc act gta gaa gaa agt ttc aag act tta ttt tgg tca     1872
Ala Ala Phe Thr Thr Val Glu Glu Ser Phe Lys Thr Leu Phe Trp Ser
    610                 615                 620

ata ttt ggg ttg tct gaa gtg act tcc gtt gtg ctc aaa tat gat cac     1920
Ile Phe Gly Leu Ser Glu Val Thr Ser Val Val Leu Lys Tyr Asp His
625                 630                 635                 640

aaa ttc ata gaa aat att gga tac gtt ctt tat gga ata tac aat gta     1968
Lys Phe Ile Glu Asn Ile Gly Tyr Val Leu Tyr Gly Ile Tyr Asn Val
                645                 650                 655

act atg gtg gtc gtt tta ctc aac atg cta att gct atg att aat agc     2016
Thr Met Val Val Val Leu Leu Asn Met Leu Ile Ala Met Ile Asn Ser
            660                 665                 670

tca tat caa gaa att gag gat gac agt gat gta gaa tgg aag ttt gct     2064
Ser Tyr Gln Glu Ile Glu Asp Asp Ser Asp Val Glu Trp Lys Phe Ala
        675                 680                 685

cgt tca aaa ctt tgg tta tcc tat ttt gat gat gga aaa aca tta cct     2112
Arg Ser Lys Leu Trp Leu Ser Tyr Phe Asp Asp Gly Lys Thr Leu Pro
    690                 695                 700

cca cct ttc agt cta gtt cct agt cca aaa tca ttt gtt tat ttc atc     2160
Pro Pro Phe Ser Leu Val Pro Ser Pro Lys Ser Phe Val Tyr Phe Ile
705                 710                 715                 720

atg cga att gtt aac ttt ccc aaa tgc aga agg aga agg ctt cag aag     2208
Met Arg Ile Val Asn Phe Pro Lys Cys Arg Arg Arg Arg Leu Gln Lys
                725                 730                 735

gat ata gaa atg gga atg ggt aac tca aag tcc agg tta aac ctc ttc     2256
Asp Ile Glu Met Gly Met Gly Asn Ser Lys Ser Arg Leu Asn Leu Phe
            740                 745                 750

act cag tct aac tca aga gtt ttt gaa tca cac agt ttt aac agc att     2304
Thr Gln Ser Asn Ser Arg Val Phe Glu Ser His Ser Phe Asn Ser Ile
        755                 760                 765

ctc aat cag cca aca cgt tat cag cag ata atg aaa aga ctt ata aag     2352
Leu Asn Gln Pro Thr Arg Tyr Gln Gln Ile Met Lys Arg Leu Ile Lys
    770                 775                 780

cgg tat gtt ttg aaa gca caa gta gac aaa gaa aat gat gaa gtt aat     2400
Arg Tyr Val Leu Lys Ala Gln Val Asp Lys Glu Asn Asp Glu Val Asn
785                 790                 795                 800

gaa ggt gaa tta aaa gaa atc aag caa gat atc tcc agc ctt cgt tat     2448
Glu Gly Glu Leu Lys Glu Ile Lys Gln Asp Ile Ser Ser Leu Arg Tyr
                805                 810                 815

gaa ctt ttg gaa gac aag agc caa gca act gag gaa tta gcc att cta     2496
Glu Leu Leu Glu Asp Lys Ser Gln Ala Thr Glu Glu Leu Ala Ile Leu
            820                 825                 830

att cat aaa ctt agt gag aaa ctg aat ccc agc atg ctg aga tgt gaa     2544
```

-continued

```
Ile His Lys Leu Ser Glu Lys Leu Asn Pro Ser Met Leu Arg Cys Glu
        835                 840                 845

tga                                                                 2547
*


<210> SEQ ID NO 30
<211> LENGTH: 250
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (68)...(89)
<223> OTHER INFORMATION: Xaa = Any amino Acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (216)...(224)
<223> OTHER INFORMATION: Xaa = Any amino Acid

<400> SEQUENCE: 30

Ile Leu Phe Ile Leu Asp Leu Leu Phe Val Leu Leu Phe Leu Leu Glu
 1               5                  10                  15

Ile Val Leu Lys Phe Ile Ala Tyr Gly Leu Lys Ser Thr Ser Asn Ile
            20                  25                  30

Ala Ala Lys Tyr Leu Lys Ser Ile Phe Asn Ile Leu Asp Leu Leu Ala
        35                  40                  45

Ile Leu Pro Leu Leu Leu Leu Leu Val Leu Phe Leu Ser Gly Thr Glu
    50                  55                  60

Gln Val Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
65                  70                  75                  80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Glu Arg Xaa Phe Xaa Ser
                85                  90                  95

Leu Glu Leu Ser Gln Tyr Arg Ile Leu Arg Phe Leu Arg Leu Leu Arg
            100                 105                 110

Leu Leu Arg Leu Leu Arg Leu Leu Arg Leu Leu Arg Arg Leu Glu Thr
        115                 120                 125

Leu Phe Glu Phe Glu Leu Gly Thr Leu Ala Trp Ser Leu Gln Ser Leu
    130                 135                 140

Gly Arg Ala Leu Lys Ser Ile Leu Arg Phe Leu Leu Leu Leu Leu Leu
145                 150                 155                 160

Leu Leu Ile Gly Phe Ser Val Ile Gly Tyr Leu Leu Phe Lys Gly Tyr
                165                 170                 175

Glu Asp Leu Ser Glu Asn Glu Val Asp Gly Asn Ser Glu Phe Ser Ser
            180                 185                 190

Tyr Phe Asp Ala Phe Tyr Phe Leu Phe Val Thr Leu Thr Thr Val Gly
        195                 200                 205

Phe Gly Asp Leu Val Pro Val Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
    210                 215                 220

Trp Leu Gly Ile Ile Phe Phe Val Leu Phe Phe Ile Ile Val Gly Leu
225                 230                 235                 240

Leu Leu Leu Asn Leu Leu Ile Ala Val Ile
                245                 250


<210> SEQ ID NO 31
<211> LENGTH: 33
<212> TYPE: PRT
```

-continued

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 31

Asp Gly Arg Thr Pro Leu His Leu Ala Ala Arg Asn Gly His Leu Glu
1 5 10 15

Val Val Lys Leu Leu Leu Glu Ala Gly Ala Asp Val Asn Ala Arg Asp
20 25 30

Lys

<210> SEQ ID NO 32
<211> LENGTH: 3220
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (338)...(2476)

<400> SEQUENCE: 32

tcgggaagcg cgccattgtg ttggtacccg ggaattcgcg gccgcgtcga ccccccctgcg 60

gagttgtgtt ttctgggaat cagcaacaaa ttgcaaagaa atggctcaaa agcttcagct 120

ctttctgtgc cctgggagct gagatgcacg tcagtggcct tgccagcgtg gccaattctc 180

tgctgactgc cagaaaaaag aggccaggaa gaaagaggaa agagaagaga tcgctcaggg 240

tctgtggtgt gtggtccatc ctcttgctga gcacattgaa aggaactggc tatctttgat 300

ctcttcctcc agatcagagt caaggaatgt gtttata atg gac act tca tcc aaa 355
Met Asp Thr Ser Ser Lys
1 5

gaa aat atc cag ttg ttc tgc aaa act tca gtg caa cct gtt gga agg 403
Glu Asn Ile Gln Leu Phe Cys Lys Thr Ser Val Gln Pro Val Gly Arg
10 15 20

cct tct ttt aaa aca gaa tat ccc tcc tca gaa gaa aag caa cca tgc 451
Pro Ser Phe Lys Thr Glu Tyr Pro Ser Ser Glu Glu Lys Gln Pro Cys
25 30 35

tgt ggt gaa cta aag gtg ttc ttg tgt gcc ttg tct ttt gtt tac ttt 499
Cys Gly Glu Leu Lys Val Phe Leu Cys Ala Leu Ser Phe Val Tyr Phe
40 45 50

gcc aaa gca ttg gca gaa ggc tat ctg aag agc acc atc act cag ata 547
Ala Lys Ala Leu Ala Glu Gly Tyr Leu Lys Ser Thr Ile Thr Gln Ile
55 60 65 70

gag aga agg ttt gat atc cct tct tca ctg gtg gga gtt att ggt ggt 595
Glu Arg Arg Phe Asp Ile Pro Ser Ser Leu Val Gly Val Ile Gly Gly
75 80 85

agt ttt gaa att ggg aat ctc tta gtt ata aca ttt gtt agc tac ttt 643
Ser Phe Glu Ile Gly Asn Leu Leu Val Ile Thr Phe Val Ser Tyr Phe
90 95 100

gga gcc aaa ctt cac agg cca aaa ata att gga gca ggg tgt gta atc 691
Gly Ala Lys Leu His Arg Pro Lys Ile Ile Gly Ala Gly Cys Val Ile
105 110 115

atg gga gtt gga aca ctg ctc att gca atg cct cag ttc ttc atg gag 739
Met Gly Val Gly Thr Leu Leu Ile Ala Met Pro Gln Phe Phe Met Glu
120 125 130

cag tac aaa tat gag aga tat tct cct tcc tcc aat tct act ctc agc 787
Gln Tyr Lys Tyr Glu Arg Tyr Ser Pro Ser Ser Asn Ser Thr Leu Ser
135 140 145 150

atc tct ccg cgt ctc cta gag tca agc agt caa tta cca gtt tca gtt 835

-continued

```
Ile Ser Pro Arg Leu Leu Glu Ser Ser Ser Gln Leu Pro Val Ser Val
                155                 160                 165

atg gaa aaa tca aaa tcc aaa ata agt aac gaa tgt gaa gtg gac act      883
Met Glu Lys Ser Lys Ser Lys Ile Ser Asn Glu Cys Glu Val Asp Thr
            170                 175                 180

agc tct tcc atg tgg att tat gtt ttc ctg ggt aat ctt ctt cgt gga      931
Ser Ser Ser Met Trp Ile Tyr Val Phe Leu Gly Asn Leu Leu Arg Gly
        185                 190                 195

ata gga gaa act ccc att cag cct ttg ggc att gcc tac ctg gat gat      979
Ile Gly Glu Thr Pro Ile Gln Pro Leu Gly Ile Ala Tyr Leu Asp Asp
    200                 205                 210

ttt gcc agt gaa gac aat gca gct ttc tat att ggg tgt gtg cag acg     1027
Phe Ala Ser Glu Asp Asn Ala Ala Phe Tyr Ile Gly Cys Val Gln Thr
215                 220                 225                 230

gtt gca att ata gga cca atc ttt ggt ttc ctg tta ggc tca tta tgt     1075
Val Ala Ile Ile Gly Pro Ile Phe Gly Phe Leu Leu Gly Ser Leu Cys
                235                 240                 245

gcc aaa cta tat gtt gac att ggc ttt gta aac cta gat cac ata acc     1123
Ala Lys Leu Tyr Val Asp Ile Gly Phe Val Asn Leu Asp His Ile Thr
            250                 255                 260

att aca cca aaa gat ccc cag tgg gta gga gcc tgg tgg ctt ggc tat     1171
Ile Thr Pro Lys Asp Pro Gln Trp Val Gly Ala Trp Trp Leu Gly Tyr
        265                 270                 275

cta ata gca gga atc ata agt ctt ctt gca gct gtg cct ttc tgg tat     1219
Leu Ile Ala Gly Ile Ile Ser Leu Leu Ala Ala Val Pro Phe Trp Tyr
    280                 285                 290

tta cca aag agt tta cca aga tcc caa agt aga gag gat tct aat tct     1267
Leu Pro Lys Ser Leu Pro Arg Ser Gln Ser Arg Glu Asp Ser Asn Ser
295                 300                 305                 310

tcc tct gag aaa tcc aag ttt att ata gat gat cac aca gac tac caa     1315
Ser Ser Glu Lys Ser Lys Phe Ile Ile Asp Asp His Thr Asp Tyr Gln
                315                 320                 325

aca ccc cag gga gaa aat gca aaa ata atg gaa atg gca aga gat ttt     1363
Thr Pro Gln Gly Glu Asn Ala Lys Ile Met Glu Met Ala Arg Asp Phe
            330                 335                 340

ctt cca tca ctg aag aat ctt ttt gga aac cca gta tac ttc cta tat     1411
Leu Pro Ser Leu Lys Asn Leu Phe Gly Asn Pro Val Tyr Phe Leu Tyr
        345                 350                 355

tta tgt aca agc act gtt cag ttc aat tct ctg ttc ggc atg gtg acg     1459
Leu Cys Thr Ser Thr Val Gln Phe Asn Ser Leu Phe Gly Met Val Thr
    360                 365                 370

tac aaa cca aag tac att gag cag cag tat gga cag tca tcc tcc agg     1507
Tyr Lys Pro Lys Tyr Ile Glu Gln Gln Tyr Gly Gln Ser Ser Ser Arg
375                 380                 385                 390

gcc aac ttt gtg atc ggg ctc atc aac att cca gca gtg gcc ctt gga     1555
Ala Asn Phe Val Ile Gly Leu Ile Asn Ile Pro Ala Val Ala Leu Gly
                395                 400                 405

ata ttc tct ggg ggg ata gtt atg aaa aaa ttc aga atc agt gtg tgt     1603
Ile Phe Ser Gly Gly Ile Val Met Lys Lys Phe Arg Ile Ser Val Cys
            410                 415                 420

gga gct gca aaa ctc tac ttg gga tca tct gtc ttt ggt tac ctc cta     1651
Gly Ala Ala Lys Leu Tyr Leu Gly Ser Ser Val Phe Gly Tyr Leu Leu
        425                 430                 435

ttt ctt tcc ctg ttt gca ctg ggc tgt gaa aat tct gat gtg gca gga     1699
Phe Leu Ser Leu Phe Ala Leu Gly Cys Glu Asn Ser Asp Val Ala Gly
    440                 445                 450

cta act gtc tcc tac caa gga acc aaa cct gtc tct tat cat gaa cga     1747
```

-continued

```
Leu Thr Val Ser Tyr Gln Gly Thr Lys Pro Val Ser Tyr His Glu Arg
455                 460                 465                 470

gct ctc ttt tca gat tgc aac tca aga tgc aaa tgt tca gag aca aaa     1795
Ala Leu Phe Ser Asp Cys Asn Ser Arg Cys Lys Cys Ser Glu Thr Lys
                475                 480                 485

tgg gaa ccc atg tgc ggt gaa aat gga atc aca tat gta tca gct tgt     1843
Trp Glu Pro Met Cys Gly Glu Asn Gly Ile Thr Tyr Val Ser Ala Cys
            490                 495                 500

ctt gct ggt tgt caa acc tcc aac agg agt gga aaa aat att ata ttt     1891
Leu Ala Gly Cys Gln Thr Ser Asn Arg Ser Gly Lys Asn Ile Ile Phe
        505                 510                 515

tac aac tgc act tgt gtg gga att gca gct tct aaa tcc gga aat tcc     1939
Tyr Asn Cys Thr Cys Val Gly Ile Ala Ala Ser Lys Ser Gly Asn Ser
    520                 525                 530

tca ggc ata gtg gga aga tgt cag aaa gac aat gga tgt ccc caa atg     1987
Ser Gly Ile Val Gly Arg Cys Gln Lys Asp Asn Gly Cys Pro Gln Met
535                 540                 545                 550

ttt ctg tat ttc ctt gta att tca gtc atc aca tcc tat act tta tcc     2035
Phe Leu Tyr Phe Leu Val Ile Ser Val Ile Thr Ser Tyr Thr Leu Ser
                555                 560                 565

cta ggt ggc ata cct gga tac ata tta ctt ctg agg tgc att aag cca     2083
Leu Gly Gly Ile Pro Gly Tyr Ile Leu Leu Leu Arg Cys Ile Lys Pro
            570                 575                 580

cag ctt aag tct ttt gcc ttg ggt atc tac aca tta gca ata aga gtt     2131
Gln Leu Lys Ser Phe Ala Leu Gly Ile Tyr Thr Leu Ala Ile Arg Val
        585                 590                 595

ctt gca gga atc cca gct cca gtg tat ttt gga gtt ttg att gat act     2179
Leu Ala Gly Ile Pro Ala Pro Val Tyr Phe Gly Val Leu Ile Asp Thr
    600                 605                 610

tca tgc ctc aaa tgg gga ttt aaa aga tgt gga agt aga gga tca tgc     2227
Ser Cys Leu Lys Trp Gly Phe Lys Arg Cys Gly Ser Arg Gly Ser Cys
615                 620                 625                 630

aga tta tat gat tca aat gtc ttc aga cat ata tat ctg gga cta act     2275
Arg Leu Tyr Asp Ser Asn Val Phe Arg His Ile Tyr Leu Gly Leu Thr
                635                 640                 645

gtg ata ctg ggc aca gtg tca att ctc cta agc att gca gta ctt ttc     2323
Val Ile Leu Gly Thr Val Ser Ile Leu Leu Ser Ile Ala Val Leu Phe
            650                 655                 660

att tta aag aaa aat tat gtt tca aaa cac aga agt ttt ata acc aag     2371
Ile Leu Lys Lys Asn Tyr Val Ser Lys His Arg Ser Phe Ile Thr Lys
        665                 670                 675

aga gaa aga aca atg gtg tct aca aga ttc caa aag gaa aat tac act     2419
Arg Glu Arg Thr Met Val Ser Thr Arg Phe Gln Lys Glu Asn Tyr Thr
    680                 685                 690

aca agt gat cat ctg cta caa ccc aac tac tgg cca ggc aag gaa act     2467
Thr Ser Asp His Leu Leu Gln Pro Asn Tyr Trp Pro Gly Lys Glu Thr
695                 700                 705                 710

caa ctt tag aaacatgatg actggaagtc atgtcttcta attggttgac             2516
Gln Leu  *

attttgcaaa caaataaatt gtaatcaaaa gagctctaaa tttgtaattt ctttctcctt   2576

tcaaaaaatg tctactttgt tttggtccta ggcattaggt aatataactg ataatatact   2636

gaaacatata atggaagatg cagatgataa aactaatttt gaacttttta atttatataa   2696

attattttat atcacttact tatttcactt tattttgctt tgtgctcatt gatatatatt   2756

agctgtactc ctagaagaac aattgtctct attgtcacac atggttatat ttaaagtaat   2816
```

-continued

```
ttctgaactg tgtaatgtgt ctagagtaag caaatactgc taacaattaa ctcatacctt   2876

gggttccttc aagtattact cctatagtat tttctcccat agctgtcttc atctgtgtat   2936

tttaataatg atcttaggat ggagcagaac atggagagga agatttcatt ttaagctcct   2996

ccttttcttt gaaatacaat aatttatata gaaatgtgta gcagcaaatt atattgggga   3056

ttagaatttt gaattaatag ctctcctact attaatttac atgtgctttt tgtgtggcgc   3116

tataagtgac tatggttgta aagtaataaa attgatgtta acatgcccaa aaaaaaaaaa   3176

aaaaaaaacc aaaaaaaaaa aaaaaaaagg gcgggccgct agac                    3220


<210> SEQ ID NO 33
<211> LENGTH: 712
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 33

Met Asp Thr Ser Ser Lys Glu Asn Ile Gln Leu Phe Cys Lys Thr Ser
 1               5                  10                  15

Val Gln Pro Val Gly Arg Pro Ser Phe Lys Thr Glu Tyr Pro Ser Ser
            20                  25                  30

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

Leu Ser Phe Val Tyr Phe Ala Lys Ala Leu Ala Glu Gly Tyr Leu Lys
    50                  55                  60

Ser Thr Ile Thr Gln Ile Glu Arg Arg Phe Asp Ile Pro Ser Ser Leu
65                  70                  75                  80

Val Gly Val Ile Gly Gly Ser Phe Glu Ile Gly Asn Leu Leu Val Ile
                85                  90                  95

Thr Phe Val Ser Tyr Phe Gly Ala Lys Leu His Arg Pro Lys Ile Ile
            100                 105                 110

Gly Ala Gly Cys Val Ile Met Gly Val Gly Thr Leu Leu Ile Ala Met
        115                 120                 125

Pro Gln Phe Phe Met Glu Gln Tyr Lys Tyr Glu Arg Tyr Ser Pro Ser
    130                 135                 140

Ser Asn Ser Thr Leu Ser Ile Ser Pro Arg Leu Leu Glu Ser Ser Ser
145                 150                 155                 160

Gln Leu Pro Val Ser Val Met Glu Lys Ser Lys Ser Lys Ile Ser Asn
                165                 170                 175

Glu Cys Glu Val Asp Thr Ser Ser Ser Met Trp Ile Tyr Val Phe Leu
            180                 185                 190

Gly Asn Leu Leu Arg Gly Ile Gly Glu Thr Pro Ile Gln Pro Leu Gly
        195                 200                 205

Ile Ala Tyr Leu Asp Asp Phe Ala Ser Glu Asp Asn Ala Ala Phe Tyr
    210                 215                 220

Ile Gly Cys Val Gln Thr Val Ala Ile Ile Gly Pro Ile Phe Gly Phe
225                 230                 235                 240

Leu Leu Gly Ser Leu Cys Ala Lys Leu Tyr Val Asp Ile Gly Phe Val
                245                 250                 255

Asn Leu Asp His Ile Thr Ile Thr Pro Lys Asp Pro Gln Trp Val Gly
            260                 265                 270

Ala Trp Trp Leu Gly Tyr Leu Ile Ala Gly Ile Ile Ser Leu Leu Ala
        275                 280                 285
```

-continued

```
Ala Val Pro Phe Trp Tyr Leu Pro Lys Ser Leu Pro Arg Ser Gln Ser
    290                 295                 300

Arg Glu Asp Ser Asn Ser Ser Ser Glu Lys Ser Lys Phe Ile Ile Asp
305                 310                 315                 320

Asp His Thr Asp Tyr Gln Thr Pro Gln Gly Glu Asn Ala Lys Ile Met
                325                 330                 335

Glu Met Ala Arg Asp Phe Leu Pro Ser Leu Lys Asn Leu Phe Gly Asn
            340                 345                 350

Pro Val Tyr Phe Leu Tyr Leu Cys Thr Ser Thr Val Gln Phe Asn Ser
        355                 360                 365

Leu Phe Gly Met Val Thr Tyr Lys Pro Lys Tyr Ile Glu Gln Gln Tyr
    370                 375                 380

Gly Gln Ser Ser Ser Arg Ala Asn Phe Val Ile Gly Leu Ile Asn Ile
385                 390                 395                 400

Pro Ala Val Ala Leu Gly Ile Phe Ser Gly Gly Ile Val Met Lys Lys
                405                 410                 415

Phe Arg Ile Ser Val Cys Gly Ala Ala Lys Leu Tyr Leu Gly Ser Ser
            420                 425                 430

Val Phe Gly Tyr Leu Leu Phe Leu Ser Leu Phe Ala Leu Gly Cys Glu
        435                 440                 445

Asn Ser Asp Val Ala Gly Leu Thr Val Ser Tyr Gln Gly Thr Lys Pro
    450                 455                 460

Val Ser Tyr His Glu Arg Ala Leu Phe Ser Asp Cys Asn Ser Arg Cys
465                 470                 475                 480

Lys Cys Ser Glu Thr Lys Trp Glu Pro Met Cys Gly Glu Asn Gly Ile
                485                 490                 495

Thr Tyr Val Ser Ala Cys Leu Ala Gly Cys Gln Thr Ser Asn Arg Ser
            500                 505                 510

Gly Lys Asn Ile Ile Phe Tyr Asn Cys Thr Cys Val Gly Ile Ala Ala
        515                 520                 525

Ser Lys Ser Gly Asn Ser Ser Gly Ile Val Gly Arg Cys Gln Lys Asp
    530                 535                 540

Asn Gly Cys Pro Gln Met Phe Leu Tyr Phe Leu Val Ile Ser Val Ile
545                 550                 555                 560

Thr Ser Tyr Thr Leu Ser Leu Gly Gly Ile Pro Gly Tyr Ile Leu Leu
                565                 570                 575

Leu Arg Cys Ile Lys Pro Gln Leu Lys Ser Phe Ala Leu Gly Ile Tyr
            580                 585                 590

Thr Leu Ala Ile Arg Val Leu Ala Gly Ile Pro Ala Pro Val Tyr Phe
        595                 600                 605

Gly Val Leu Ile Asp Thr Ser Cys Leu Lys Trp Gly Phe Lys Arg Cys
    610                 615                 620

Gly Ser Arg Gly Ser Cys Arg Leu Tyr Asp Ser Asn Val Phe Arg His
625                 630                 635                 640

Ile Tyr Leu Gly Leu Thr Val Ile Leu Gly Thr Val Ser Ile Leu Leu
                645                 650                 655

Ser Ile Ala Val Leu Phe Ile Leu Lys Lys Asn Tyr Val Ser Lys His
            660                 665                 670

Arg Ser Phe Ile Thr Lys Arg Glu Arg Thr Met Val Ser Thr Arg Phe
        675                 680                 685

Gln Lys Glu Asn Tyr Thr Thr Ser Asp His Leu Leu Gln Pro Asn Tyr
```

-continued

```
    690                 695                 700

Trp Pro Gly Lys Glu Thr Gln Leu
705                 710


<210> SEQ ID NO 34
<211> LENGTH: 2139
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(2139)

<400> SEQUENCE: 34

atg gac act tca tcc aaa gaa aat atc cag ttg ttc tgc aaa act tca       48
Met Asp Thr Ser Ser Lys Glu Asn Ile Gln Leu Phe Cys Lys Thr Ser
 1               5                  10                  15

gtg caa cct gtt gga agg cct tct ttt aaa aca gaa tat ccc tcc tca       96
Val Gln Pro Val Gly Arg Pro Ser Phe Lys Thr Glu Tyr Pro Ser Ser
            20                  25                  30

gaa gaa aag caa cca tgc tgt ggt gaa cta aag gtg ttc ttg tgt gcc      144
Glu Glu Lys Gln Pro Cys Cys Gly Glu Leu Lys Val Phe Leu Cys Ala
        35                  40                  45

ttg tct ttt gtt tac ttt gcc aaa gca ttg gca gaa ggc tat ctg aag      192
Leu Ser Phe Val Tyr Phe Ala Lys Ala Leu Ala Glu Gly Tyr Leu Lys
    50                  55                  60

agc acc atc act cag ata gag aga agg ttt gat atc cct tct tca ctg      240
Ser Thr Ile Thr Gln Ile Glu Arg Arg Phe Asp Ile Pro Ser Ser Leu
 65                  70                  75                  80

gtg gga gtt att ggt ggt agt ttt gaa att ggg aat ctc tta gtt ata      288
Val Gly Val Ile Gly Gly Ser Phe Glu Ile Gly Asn Leu Leu Val Ile
                85                  90                  95

aca ttt gtt agc tac ttt gga gcc aaa ctt cac agg cca aaa ata att      336
Thr Phe Val Ser Tyr Phe Gly Ala Lys Leu His Arg Pro Lys Ile Ile
            100                 105                 110

gga gca ggg tgt gta atc atg gga gtt gga aca ctg ctc att gca atg      384
Gly Ala Gly Cys Val Ile Met Gly Val Gly Thr Leu Leu Ile Ala Met
        115                 120                 125

cct cag ttc ttc atg gag cag tac aaa tat gag aga tat tct cct tcc      432
Pro Gln Phe Phe Met Glu Gln Tyr Lys Tyr Glu Arg Tyr Ser Pro Ser
    130                 135                 140

tcc aat tct act ctc agc atc tct ccg cgt ctc cta gag tca agc agt      480
Ser Asn Ser Thr Leu Ser Ile Ser Pro Arg Leu Leu Glu Ser Ser Ser
145                 150                 155                 160

caa tta cca gtt tca gtt atg gaa aaa tca aaa tcc aaa ata agt aac      528
Gln Leu Pro Val Ser Val Met Glu Lys Ser Lys Ser Lys Ile Ser Asn
                165                 170                 175

gaa tgt gaa gtg gac act agc tct tcc atg tgg att tat gtt ttc ctg      576
Glu Cys Glu Val Asp Thr Ser Ser Ser Met Trp Ile Tyr Val Phe Leu
            180                 185                 190

ggt aat ctt ctt cgt gga ata gga gaa act ccc att cag cct ttg ggc      624
Gly Asn Leu Leu Arg Gly Ile Gly Glu Thr Pro Ile Gln Pro Leu Gly
        195                 200                 205

att gcc tac ctg gat gat ttt gcc agt gaa gac aat gca gct ttc tat      672
Ile Ala Tyr Leu Asp Asp Phe Ala Ser Glu Asp Asn Ala Ala Phe Tyr
    210                 215                 220

att ggg tgt gtg cag acg gtt gca att ata gga cca atc ttt ggt ttc      720
Ile Gly Cys Val Gln Thr Val Ala Ile Ile Gly Pro Ile Phe Gly Phe
225                 230                 235                 240
```

-continued

```
ctg tta ggc tca tta tgt gcc aaa cta tat gtt gac att ggc ttt gta      768
Leu Leu Gly Ser Leu Cys Ala Lys Leu Tyr Val Asp Ile Gly Phe Val
                245                 250                 255

aac cta gat cac ata acc att aca cca aaa gat ccc cag tgg gta gga      816
Asn Leu Asp His Ile Thr Ile Thr Pro Lys Asp Pro Gln Trp Val Gly
            260                 265                 270

gcc tgg tgg ctt ggc tat cta ata gca gga atc ata agt ctt ctt gca      864
Ala Trp Trp Leu Gly Tyr Leu Ile Ala Gly Ile Ile Ser Leu Leu Ala
        275                 280                 285

gct gtg cct ttc tgg tat tta cca aag agt tta cca aga tcc caa agt      912
Ala Val Pro Phe Trp Tyr Leu Pro Lys Ser Leu Pro Arg Ser Gln Ser
    290                 295                 300

aga gag gat tct aat tct tcc tct gag aaa tcc aag ttt att ata gat      960
Arg Glu Asp Ser Asn Ser Ser Ser Glu Lys Ser Lys Phe Ile Ile Asp
305                 310                 315                 320

gat cac aca gac tac caa aca ccc cag gga gaa aat gca aaa ata atg     1008
Asp His Thr Asp Tyr Gln Thr Pro Gln Gly Glu Asn Ala Lys Ile Met
                325                 330                 335

gaa atg gca aga gat ttt ctt cca tca ctg aag aat ctt ttt gga aac     1056
Glu Met Ala Arg Asp Phe Leu Pro Ser Leu Lys Asn Leu Phe Gly Asn
            340                 345                 350

cca gta tac ttc cta tat tta tgt aca agc act gtt cag ttc aat tct     1104
Pro Val Tyr Phe Leu Tyr Leu Cys Thr Ser Thr Val Gln Phe Asn Ser
        355                 360                 365

ctg ttc ggc atg gtg acg tac aaa cca aag tac att gag cag cag tat     1152
Leu Phe Gly Met Val Thr Tyr Lys Pro Lys Tyr Ile Glu Gln Gln Tyr
    370                 375                 380

gga cag tca tcc tcc agg gcc aac ttt gtg atc ggg ctc atc aac att     1200
Gly Gln Ser Ser Ser Arg Ala Asn Phe Val Ile Gly Leu Ile Asn Ile
385                 390                 395                 400

cca gca gtg gcc ctt gga ata ttc tct ggg ggg ata gtt atg aaa aaa     1248
Pro Ala Val Ala Leu Gly Ile Phe Ser Gly Gly Ile Val Met Lys Lys
                405                 410                 415

ttc aga atc agt gtg tgt gga gct gca aaa ctc tac ttg gga tca tct     1296
Phe Arg Ile Ser Val Cys Gly Ala Ala Lys Leu Tyr Leu Gly Ser Ser
            420                 425                 430

gtc ttt ggt tac ctc cta ttt ctt tcc ctg ttt gca ctg ggc tgt gaa     1344
Val Phe Gly Tyr Leu Leu Phe Leu Ser Leu Phe Ala Leu Gly Cys Glu
        435                 440                 445

aat tct gat gtg gca gga cta act gtc tcc tac caa gga acc aaa cct     1392
Asn Ser Asp Val Ala Gly Leu Thr Val Ser Tyr Gln Gly Thr Lys Pro
    450                 455                 460

gtc tct tat cat gaa cga gct ctc ttt tca gat tgc aac tca aga tgc     1440
Val Ser Tyr His Glu Arg Ala Leu Phe Ser Asp Cys Asn Ser Arg Cys
465                 470                 475                 480

aaa tgt tca gag aca aaa tgg gaa ccc atg tgc ggt gaa aat gga atc     1488
Lys Cys Ser Glu Thr Lys Trp Glu Pro Met Cys Gly Glu Asn Gly Ile
                485                 490                 495

aca tat gta tca gct tgt ctt gct ggt tgt caa acc tcc aac agg agt     1536
Thr Tyr Val Ser Ala Cys Leu Ala Gly Cys Gln Thr Ser Asn Arg Ser
            500                 505                 510

gga aaa aat att ata ttt tac aac tgc act tgt gtg gga att gca gct     1584
Gly Lys Asn Ile Ile Phe Tyr Asn Cys Thr Cys Val Gly Ile Ala Ala
        515                 520                 525

tct aaa tcc gga aat tcc tca ggc ata gtg gga aga tgt cag aaa gac     1632
Ser Lys Ser Gly Asn Ser Ser Gly Ile Val Gly Arg Cys Gln Lys Asp
    530                 535                 540
```

-continued

```
aat gga tgt ccc caa atg ttt ctg tat ttc ctt gta att tca gtc atc     1680
Asn Gly Cys Pro Gln Met Phe Leu Tyr Phe Leu Val Ile Ser Val Ile
545                 550                 555                 560

aca tcc tat act tta tcc cta ggt ggc ata cct gga tac ata tta ctt     1728
Thr Ser Tyr Thr Leu Ser Leu Gly Gly Ile Pro Gly Tyr Ile Leu Leu
                565                 570                 575

ctg agg tgc att aag cca cag ctt aag tct ttt gcc ttg ggt atc tac     1776
Leu Arg Cys Ile Lys Pro Gln Leu Lys Ser Phe Ala Leu Gly Ile Tyr
            580                 585                 590

aca tta gca ata aga gtt ctt gca gga atc cca gct cca gtg tat ttt     1824
Thr Leu Ala Ile Arg Val Leu Ala Gly Ile Pro Ala Pro Val Tyr Phe
        595                 600                 605

gga gtt ttg att gat act tca tgc ctc aaa tgg gga ttt aaa aga tgt     1872
Gly Val Leu Ile Asp Thr Ser Cys Leu Lys Trp Gly Phe Lys Arg Cys
    610                 615                 620

gga agt aga gga tca tgc aga tta tat gat tca aat gtc ttc aga cat     1920
Gly Ser Arg Gly Ser Cys Arg Leu Tyr Asp Ser Asn Val Phe Arg His
625                 630                 635                 640

ata tat ctg gga cta act gtg ata ctg ggc aca gtg tca att ctc cta     1968
Ile Tyr Leu Gly Leu Thr Val Ile Leu Gly Thr Val Ser Ile Leu Leu
                645                 650                 655

agc att gca gta ctt ttc att tta aag aaa aat tat gtt tca aaa cac     2016
Ser Ile Ala Val Leu Phe Ile Leu Lys Lys Asn Tyr Val Ser Lys His
            660                 665                 670

aga agt ttt ata acc aag aga gaa aga aca atg gtg tct aca aga ttc     2064
Arg Ser Phe Ile Thr Lys Arg Glu Arg Thr Met Val Ser Thr Arg Phe
        675                 680                 685

caa aag gaa aat tac act aca agt gat cat ctg cta caa ccc aac tac     2112
Gln Lys Glu Asn Tyr Thr Thr Ser Asp His Leu Leu Gln Pro Asn Tyr
    690                 695                 700

tgg cca ggc aag gaa act caa ctt tag                                 2139
Trp Pro Gly Lys Glu Thr Gln Leu  *
705                 710


<210> SEQ ID NO 35
<211> LENGTH: 2866
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (365)...(2026)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(2866)
<223> OTHER INFORMATION: n = A,T,C or G

<400> SEQUENCE: 35

aagtgtcacc ggtcacactt taattccagt ctaaaattaa agtcttcagt ctccacattc     60

cctactttcc aaattcagct ttcccgggag gtctggagca gctgcctctc tggggagatg    120

ctggaggtct cggaatcacc tcacacggcc tcagggccca gttggagcca ccccaagtga    180

caccagcagg cagatgacca gagagcctga gcctccggcc ccgagtctgt gaagcctagc    240

cgctgggctg gagaagccac tgtgggcacc accgtggggg aaacaggccc gttgccctgg    300

cctctttgcc ctgggccagc ctttgtgaag tgggcccctc ttctgggccc cttgagtagg    360

ttcc atg gca ttt tct gaa ctc ctg gac ctc gtg ggt ggc ctg ggc agg     409
Met Ala Phe Ser Glu Leu Leu Asp Leu Val Gly Gly Leu Gly Arg
 1               5                  10                  15

ttc cag gtt ctc cag acg atg gct ctg atg gtc tcc atc atg tgg ctg      457
```

-continued

```
Phe Gln Val Leu Gln Thr Met Ala Leu Met Val Ser Ile Met Trp Leu
                20                  25                  30

tgt acc cag agc atg ctg gag aac ttc tcg gcc gcc gtg ccc agc cac      505
Cys Thr Gln Ser Met Leu Glu Asn Phe Ser Ala Ala Val Pro Ser His
            35                  40                  45

cgc tgc tgg gca ccc ctc ctg gac aac agc acg gct cag gcc agc atc      553
Arg Cys Trp Ala Pro Leu Leu Asp Asn Ser Thr Ala Gln Ala Ser Ile
        50                  55                  60

cta ggg agc ttg agt cct gag gcc ctc ctg gct att tcc atc ccg ccg      601
Leu Gly Ser Leu Ser Pro Glu Ala Leu Leu Ala Ile Ser Ile Pro Pro
    65                  70                  75

ggc ccc aac cag agg ccc cac cag tgc cgc cgc ttc cgc cag cca cag      649
Gly Pro Asn Gln Arg Pro His Gln Cys Arg Arg Phe Arg Gln Pro Gln
80                  85                  90                  95

tgg cag ctc ttg gac ccc aat gcc acg gcc acc agc tgg agc gag gcc      697
Trp Gln Leu Leu Asp Pro Asn Ala Thr Ala Thr Ser Trp Ser Glu Ala
                100                 105                 110

gac acg gag ccg tgt gtg gat ggc tgg gtc tat gac cgc agc atc ttc      745
Asp Thr Glu Pro Cys Val Asp Gly Trp Val Tyr Asp Arg Ser Ile Phe
            115                 120                 125

acc tcc aca atc gtg gcc aag tgg aac ctc gtg tgt gac tct cat gct      793
Thr Ser Thr Ile Val Ala Lys Trp Asn Leu Val Cys Asp Ser His Ala
        130                 135                 140

ctg aag ccc atg gcc cag tcc atc tac ctg gct ggg att ctg gtg gga      841
Leu Lys Pro Met Ala Gln Ser Ile Tyr Leu Ala Gly Ile Leu Val Gly
    145                 150                 155

gct gct gcg tgc ggc cct gcc tca gac agg ttt ggg cgc agg ctg gtg      889
Ala Ala Ala Cys Gly Pro Ala Ser Asp Arg Phe Gly Arg Arg Leu Val
160                 165                 170                 175

cta acc tgg agc tac ctt cag atg gct gtg atg ggt acg gca gct gcc      937
Leu Thr Trp Ser Tyr Leu Gln Met Ala Val Met Gly Thr Ala Ala Ala
                180                 185                 190

ttc gcc cct gcc ttc ccc gtg tac tgn ntg ntc cgc ttc ctg ttg gcc      985
Phe Ala Pro Ala Phe Pro Val Tyr Xaa Xaa Xaa Arg Phe Leu Leu Ala
            195                 200                 205

ttc cac gcg tcc ggg gct gaa cca ctc ggt ctc ctt gca gtg atg gag     1033
Phe His Ala Ser Gly Ala Glu Pro Leu Gly Leu Leu Ala Val Met Glu
        210                 215                 220

tgg acg gcg gca cgg gcc cga ccc ttg gtg atg acc ttg aac tct ctg     1081
Trp Thr Ala Ala Arg Ala Arg Pro Leu Val Met Thr Leu Asn Ser Leu
    225                 230                 235

ggc ttc agc ttc ggc cat ggc ctg aca gct gca gtg gcc tac ggt gtg     1129
Gly Phe Ser Phe Gly His Gly Leu Thr Ala Ala Val Ala Tyr Gly Val
240                 245                 250                 255

cgg gac tgg aca ctg ctg cag ctg gtg gtc tcg gtc ccc ttc ttc ctc     1177
Arg Asp Trp Thr Leu Leu Gln Leu Val Val Ser Val Pro Phe Phe Leu
                260                 265                 270

tgc ttt ttg tac tcc tgg tgg ctg gca gag tcg gca cga tgg ctc ctc     1225
Cys Phe Leu Tyr Ser Trp Trp Leu Ala Glu Ser Ala Arg Trp Leu Leu
            275                 280                 285

acc aca ggc agg ctg gat tgg ggc ctg cag gag ctg tgg agg gtg gct     1273
Thr Thr Gly Arg Leu Asp Trp Gly Leu Gln Glu Leu Trp Arg Val Ala
        290                 295                 300

ccc atc aac gga aag ggg gca gtg cag gac acc ctg acc cct gag gtc     1321
Pro Ile Asn Gly Lys Gly Ala Val Gln Asp Thr Leu Thr Pro Glu Val
    305                 310                 315

ttg ctt tca gcc atg cgg gag gag ctg agc atg ggc cag cct cct gcc     1369
```

-continued

```
Leu Leu Ser Ala Met Arg Glu Glu Leu Ser Met Gly Gln Pro Pro Ala
320                 325                 330                 335

agc ctg ggc acc ctg ctc cgc atg ccc gga ctg cgc ttc cgg acc tgt     1417
Ser Leu Gly Thr Leu Leu Arg Met Pro Gly Leu Arg Phe Arg Thr Cys
                340                 345                 350

atc tcc acg ttg tgc tgg ttc gcc ttt ggc ttc acc ttc ttc ggc ctg     1465
Ile Ser Thr Leu Cys Trp Phe Ala Phe Gly Phe Thr Phe Phe Gly Leu
            355                 360                 365

gcc ctg gac ctg cag gcc ctg ggc agc aac atc ttc ctg ctc caa atg     1513
Ala Leu Asp Leu Gln Ala Leu Gly Ser Asn Ile Phe Leu Leu Gln Met
        370                 375                 380

ttc att ggt gtc gtg gac atc cca gcc aag atg ggc gcc ctg ctg ctg     1561
Phe Ile Gly Val Val Asp Ile Pro Ala Lys Met Gly Ala Leu Leu Leu
    385                 390                 395

ctg agc cac ctg ggc cgc cgc ccc acg ctg gcc gca tcc ctg ttg ctg     1609
Leu Ser His Leu Gly Arg Arg Pro Thr Leu Ala Ala Ser Leu Leu Leu
400                 405                 410                 415

gcg ggg ctc tgc att ctg gcc aac acg ctg gtg ccc cac gaa atg ggg     1657
Ala Gly Leu Cys Ile Leu Ala Asn Thr Leu Val Pro His Glu Met Gly
                420                 425                 430

gct ctg cgc tca gcc ttg gcc gtg ctg ggg ctg ggc ggg gtg ggg gct     1705
Ala Leu Arg Ser Ala Leu Ala Val Leu Gly Leu Gly Gly Val Gly Ala
            435                 440                 445

gcc ttc acc tgc atc acc atc tac agc agc gag ctc ttc ccc act gtg     1753
Ala Phe Thr Cys Ile Thr Ile Tyr Ser Ser Glu Leu Phe Pro Thr Val
        450                 455                 460

ctc agg atg acg gca gtg ggc ttg ggc cag atg gca gcc cgt gga gga     1801
Leu Arg Met Thr Ala Val Gly Leu Gly Gln Met Ala Ala Arg Gly Gly
    465                 470                 475

gcc atc ctg ggg cct ctg gtc cgg ctg ctg ggt gtc cat ggc ccc tgg     1849
Ala Ile Leu Gly Pro Leu Val Arg Leu Leu Gly Val His Gly Pro Trp
480                 485                 490                 495

ctg ccc ttg ctg gtg tat ggg acg gtg cca gtg ctg agt ggc ctg gcc     1897
Leu Pro Leu Leu Val Tyr Gly Thr Val Pro Val Leu Ser Gly Leu Ala
                500                 505                 510

gca ctg ctt ctg ccc gag acc cag agc ttg ccg ctg ccc gac acc atc     1945
Ala Leu Leu Leu Pro Glu Thr Gln Ser Leu Pro Leu Pro Asp Thr Ile
            515                 520                 525

caa gat gtg cag aac cag gca gta aag aag gca aca cat ggc acg ctg     1993
Gln Asp Val Gln Asn Gln Ala Val Lys Lys Ala Thr His Gly Thr Leu
        530                 535                 540

ggg aac tct gtc cta aaa tcc aca cag ttt tag cctcctgggg aacctgcgat   2046
Gly Asn Ser Val Leu Lys Ser Thr Gln Phe  *
    545                 550

gggacggtca gaggaagaga cttcttctgt tctctggaga aggcaggagg aaagcaaaga   2106

cctccatttc cagaggccca gaggctgccc tctgaggtcc ccactctccc ccagggctgc   2166

ccctccaggt gagccctgcc cctctcacag tccaaggggc ccccttcaat actgaagggg   2226

aaaaggacag tttgattggc aggaggtgac ccagtgcacc atcaccctgc cctgccctcg   2286

tggcttcgga gagcagaggg gtcaggccca ggggaacgag ctggccttgc caaccctctg   2346

cttgactccg cactgccact tgtccccccca cacccgtcca cctgcccaga gctcagagct   2406

aaccaccatc catggtcaag acctctccta gctccacaca agcagtagag tctcagctcc   2466

acagctttac ccagaagccc tgtaagcctg gcccctggcc cctccccatg tccctccagg   2526

cctcagccac ctgcccacca catcctctgc ctgctgtccc cttcccaccc tcatccctga   2586
```

-continued

```
ccgactccac ttaaccccca aacccagccc cccttccagg ggtccagggc cagcctgaga   2646

tgcccgtgaa actcctaccc acagttacag ccacaagcct gcctcctccc accctgccag   2706

cctatgagtt cccagagggt tggggcagtc ccatgacccc atgtcccagc tccccacaca   2766

gcgctgggcc agagaggcat tggtgcgagg gattgaataa agaaacaaat gaaaaaaaaa   2826

aaaaaaaaaa agcattgcgg ccgcaagctt tttcccttta                         2866


<210> SEQ ID NO 36
<211> LENGTH: 553
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (200)...(202)
<223> OTHER INFORMATION: Xaa = Any Amino Acid

<400> SEQUENCE: 36

Met Ala Phe Ser Glu Leu Leu Asp Leu Val Gly Gly Leu Gly Arg Phe
 1               5                  10                  15

Gln Val Leu Gln Thr Met Ala Leu Met Val Ser Ile Met Trp Leu Cys
            20                  25                  30

Thr Gln Ser Met Leu Glu Asn Phe Ser Ala Ala Val Pro Ser His Arg
        35                  40                  45

Cys Trp Ala Pro Leu Leu Asp Asn Ser Thr Ala Gln Ala Ser Ile Leu
    50                  55                  60

Gly Ser Leu Ser Pro Glu Ala Leu Leu Ala Ile Ser Ile Pro Pro Gly
65                  70                  75                  80

Pro Asn Gln Arg Pro His Gln Cys Arg Arg Phe Arg Gln Pro Gln Trp
                85                  90                  95

Gln Leu Leu Asp Pro Asn Ala Thr Ala Thr Ser Trp Ser Glu Ala Asp
            100                 105                 110

Thr Glu Pro Cys Val Asp Gly Trp Val Tyr Asp Arg Ser Ile Phe Thr
        115                 120                 125

Ser Thr Ile Val Ala Lys Trp Asn Leu Val Cys Asp Ser His Ala Leu
    130                 135                 140

Lys Pro Met Ala Gln Ser Ile Tyr Leu Ala Gly Ile Leu Val Gly Ala
145                 150                 155                 160

Ala Ala Cys Gly Pro Ala Ser Asp Arg Phe Gly Arg Arg Leu Val Leu
                165                 170                 175

Thr Trp Ser Tyr Leu Gln Met Ala Val Met Gly Thr Ala Ala Ala Phe
            180                 185                 190

Ala Pro Ala Phe Pro Val Tyr Xaa Xaa Xaa Arg Phe Leu Leu Ala Phe
        195                 200                 205

His Ala Ser Gly Ala Glu Pro Leu Gly Leu Leu Ala Val Met Glu Trp
    210                 215                 220

Thr Ala Ala Arg Ala Arg Pro Leu Val Met Thr Leu Asn Ser Leu Gly
225                 230                 235                 240

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

Asp Trp Thr Leu Leu Gln Leu Val Val Ser Val Pro Phe Phe Leu Cys
            260                 265                 270

Phe Leu Tyr Ser Trp Trp Leu Ala Glu Ser Ala Arg Trp Leu Leu Thr
        275                 280                 285
```

-continued

```
Thr Gly Arg Leu Asp Trp Gly Leu Gln Glu Leu Trp Arg Val Ala Pro
    290                 295                 300

Ile Asn Gly Lys Gly Ala Val Gln Asp Thr Leu Thr Pro Glu Val Leu
305                 310                 315                 320

Leu Ser Ala Met Arg Glu Glu Leu Ser Met Gly Gln Pro Pro Ala Ser
                325                 330                 335

Leu Gly Thr Leu Leu Arg Met Pro Gly Leu Arg Phe Arg Thr Cys Ile
            340                 345                 350

Ser Thr Leu Cys Trp Phe Ala Phe Gly Phe Thr Phe Phe Gly Leu Ala
        355                 360                 365

Leu Asp Leu Gln Ala Leu Gly Ser Asn Ile Phe Leu Leu Gln Met Phe
    370                 375                 380

Ile Gly Val Val Asp Ile Pro Ala Lys Met Gly Ala Leu Leu Leu Leu
385                 390                 395                 400

Ser His Leu Gly Arg Arg Pro Thr Leu Ala Ala Ser Leu Leu Leu Ala
                405                 410                 415

Gly Leu Cys Ile Leu Ala Asn Thr Leu Val Pro His Glu Met Gly Ala
            420                 425                 430

Leu Arg Ser Ala Leu Ala Val Leu Gly Leu Gly Gly Val Gly Ala Ala
        435                 440                 445

Phe Thr Cys Ile Thr Ile Tyr Ser Ser Glu Leu Phe Pro Thr Val Leu
    450                 455                 460

Arg Met Thr Ala Val Gly Leu Gly Gln Met Ala Ala Arg Gly Gly Ala
465                 470                 475                 480

Ile Leu Gly Pro Leu Val Arg Leu Leu Gly Val His Gly Pro Trp Leu
                485                 490                 495

Pro Leu Leu Val Tyr Gly Thr Val Pro Val Leu Ser Gly Leu Ala Ala
            500                 505                 510

Leu Leu Leu Pro Glu Thr Gln Ser Leu Pro Leu Pro Asp Thr Ile Gln
        515                 520                 525

Asp Val Gln Asn Gln Ala Val Lys Lys Ala Thr His Gly Thr Leu Gly
    530                 535                 540

Asn Ser Val Leu Lys Ser Thr Gln Phe
545                 550


<210> SEQ ID NO 37
<211> LENGTH: 1662
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(1662)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(1662)
<223> OTHER INFORMATION: n = A,T,C or G

<400> SEQUENCE: 37

atg gca ttt tct gaa ctc ctg gac ctc gtg ggt ggc ctg ggc agg ttc       48
Met Ala Phe Ser Glu Leu Leu Asp Leu Val Gly Gly Leu Gly Arg Phe
 1               5                  10                  15

cag gtt ctc cag acg atg gct ctg atg gtc tcc atc atg tgg ctg tgt       96
Gln Val Leu Gln Thr Met Ala Leu Met Val Ser Ile Met Trp Leu Cys
            20                  25                  30

acc cag agc atg ctg gag aac ttc tcg gcc gcc gtg ccc agc cac cgc      144
```

-continued

```
Thr Gln Ser Met Leu Glu Asn Phe Ser Ala Ala Val Pro Ser His Arg
         35                  40                  45

tgc tgg gca ccc ctc ctg gac aac agc acg gct cag gcc agc atc cta      192
Cys Trp Ala Pro Leu Leu Asp Asn Ser Thr Ala Gln Ala Ser Ile Leu
     50                  55                  60

ggg agc ttg agt cct gag gcc ctc ctg gct att tcc atc ccg ccg ggc      240
Gly Ser Leu Ser Pro Glu Ala Leu Leu Ala Ile Ser Ile Pro Pro Gly
 65                  70                  75                  80

ccc aac cag agg ccc cac cag tgc cgc cgc ttc cgc cag cca cag tgg      288
Pro Asn Gln Arg Pro His Gln Cys Arg Arg Phe Arg Gln Pro Gln Trp
                 85                  90                  95

cag ctc ttg gac ccc aat gcc acg gcc acc agc tgg agc gag gcc gac      336
Gln Leu Leu Asp Pro Asn Ala Thr Ala Thr Ser Trp Ser Glu Ala Asp
            100                 105                 110

acg gag ccg tgt gtg gat ggc tgg gtc tat gac cgc agc atc ttc acc      384
Thr Glu Pro Cys Val Asp Gly Trp Val Tyr Asp Arg Ser Ile Phe Thr
        115                 120                 125

tcc aca atc gtg gcc aag tgg aac ctc gtg tgt gac tct cat gct ctg      432
Ser Thr Ile Val Ala Lys Trp Asn Leu Val Cys Asp Ser His Ala Leu
    130                 135                 140

aag ccc atg gcc cag tcc atc tac ctg gct ggg att ctg gtg gga gct      480
Lys Pro Met Ala Gln Ser Ile Tyr Leu Ala Gly Ile Leu Val Gly Ala
145                 150                 155                 160

gct gcg tgc ggc cct gcc tca gac agg ttt ggg cgc agg ctg gtg cta      528
Ala Ala Cys Gly Pro Ala Ser Asp Arg Phe Gly Arg Arg Leu Val Leu
                165                 170                 175

acc tgg agc tac ctt cag atg gct gtg atg ggt acg gca gct gcc ttc      576
Thr Trp Ser Tyr Leu Gln Met Ala Val Met Gly Thr Ala Ala Ala Phe
            180                 185                 190

gcc cct gcc ttc ccc gtg tac tgn ntg ntc cgc ttc ctg ttg gcc ttc      624
Ala Pro Ala Phe Pro Val Tyr Xaa Xaa Xaa Arg Phe Leu Leu Ala Phe
        195                 200                 205

cac gcg tcc ggg gct gaa cca ctc ggt ctc ctt gca gtg atg gag tgg      672
His Ala Ser Gly Ala Glu Pro Leu Gly Leu Leu Ala Val Met Glu Trp
    210                 215                 220

acg gcg gca cgg gcc cga ccc ttg gtg atg acc ttg aac tct ctg ggc      720
Thr Ala Ala Arg Ala Arg Pro Leu Val Met Thr Leu Asn Ser Leu Gly
225                 230                 235                 240

ttc agc ttc ggc cat ggc ctg aca gct gca gtg gcc tac ggt gtg cgg      768
Phe Ser Phe Gly His Gly Leu Thr Ala Ala Val Ala Tyr Gly Val Arg
                245                 250                 255

gac tgg aca ctg ctg cag ctg gtg gtc tcg gtc ccc ttc ttc ctc tgc      816
Asp Trp Thr Leu Leu Gln Leu Val Val Ser Val Pro Phe Phe Leu Cys
            260                 265                 270

ttt ttg tac tcc tgg tgg ctg gca gag tcg gca cga tgg ctc ctc acc      864
Phe Leu Tyr Ser Trp Trp Leu Ala Glu Ser Ala Arg Trp Leu Leu Thr
        275                 280                 285

aca ggc agg ctg gat tgg ggc ctg cag gag ctg tgg agg gtg gct ccc      912
Thr Gly Arg Leu Asp Trp Gly Leu Gln Glu Leu Trp Arg Val Ala Pro
    290                 295                 300

atc aac gga aag ggg gca gtg cag gac acc ctg acc cct gag gtc ttg      960
Ile Asn Gly Lys Gly Ala Val Gln Asp Thr Leu Thr Pro Glu Val Leu
305                 310                 315                 320

ctt tca gcc atg cgg gag gag ctg agc atg ggc cag cct cct gcc agc     1008
Leu Ser Ala Met Arg Glu Glu Leu Ser Met Gly Gln Pro Pro Ala Ser
                325                 330                 335

ctg ggc acc ctg ctc cgc atg ccc gga ctg cgc ttc cgg acc tgt atc     1056
```

-continued

```
Leu Gly Thr Leu Leu Arg Met Pro Gly Leu Arg Phe Arg Thr Cys Ile
            340                 345                 350

tcc acg ttg tgc tgg ttc gcc ttt ggc ttc acc ttc ttc ggc ctg gcc     1104
Ser Thr Leu Cys Trp Phe Ala Phe Gly Phe Thr Phe Phe Gly Leu Ala
        355                 360                 365

ctg gac ctg cag gcc ctg ggc agc aac atc ttc ctg ctc caa atg ttc     1152
Leu Asp Leu Gln Ala Leu Gly Ser Asn Ile Phe Leu Leu Gln Met Phe
    370                 375                 380

att ggt gtc gtg gac atc cca gcc aag atg ggc gcc ctg ctg ctg ctg     1200
Ile Gly Val Val Asp Ile Pro Ala Lys Met Gly Ala Leu Leu Leu Leu
385                 390                 395                 400

agc cac ctg ggc cgc cgc ccc acg ctg gcc gca tcc ctg ttg ctg gcg     1248
Ser His Leu Gly Arg Arg Pro Thr Leu Ala Ala Ser Leu Leu Leu Ala
                405                 410                 415

ggg ctc tgc att ctg gcc aac acg ctg gtg ccc cac gaa atg ggg gct     1296
Gly Leu Cys Ile Leu Ala Asn Thr Leu Val Pro His Glu Met Gly Ala
            420                 425                 430

ctg cgc tca gcc ttg gcc gtg ctg ggg ctg ggc ggg gtg ggg gct gcc     1344
Leu Arg Ser Ala Leu Ala Val Leu Gly Leu Gly Gly Val Gly Ala Ala
        435                 440                 445

ttc acc tgc atc acc atc tac agc agc gag ctc ttc ccc act gtg ctc     1392
Phe Thr Cys Ile Thr Ile Tyr Ser Ser Glu Leu Phe Pro Thr Val Leu
    450                 455                 460

agg atg acg gca gtg ggc ttg ggc cag atg gca gcc cgt gga gga gcc     1440
Arg Met Thr Ala Val Gly Leu Gly Gln Met Ala Ala Arg Gly Gly Ala
465                 470                 475                 480

atc ctg ggg cct ctg gtc cgg ctg ctg ggt gtc cat ggc ccc tgg ctg     1488
Ile Leu Gly Pro Leu Val Arg Leu Leu Gly Val His Gly Pro Trp Leu
                485                 490                 495

ccc ttg ctg gtg tat ggg acg gtg cca gtg ctg agt ggc ctg gcc gca     1536
Pro Leu Leu Val Tyr Gly Thr Val Pro Val Leu Ser Gly Leu Ala Ala
            500                 505                 510

ctg ctt ctg ccc gag acc cag agc ttg ccg ctg ccc gac acc atc caa     1584
Leu Leu Leu Pro Glu Thr Gln Ser Leu Pro Leu Pro Asp Thr Ile Gln
        515                 520                 525

gat gtg cag aac cag gca gta aag aag gca aca cat ggc acg ctg ggg     1632
Asp Val Gln Asn Gln Ala Val Lys Lys Ala Thr His Gly Thr Leu Gly
    530                 535                 540

aac tct gtc cta aaa tcc aca cag ttt tag                             1662
Asn Ser Val Leu Lys Ser Thr Gln Phe  *
545                 550


<210> SEQ ID NO 38
<211> LENGTH: 2480
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (148)...(2307)

<400> SEQUENCE: 38

gtcgacccac gcgtccgggc ggcaagagcc ccttgtggcc accgagtcct ccgacgccct     60

cgccaggctg gcctttgggt tggcccaggc aggacgggca gccgagagca ctcgggccgc    120

gtcgccagga gccgcccagg gtgagcc atg ttc gta ggc gtc gcc cgg cac tct    174
Met Phe Val Gly Val Ala Arg His Ser
 1               5

ggg agc cag gat gaa gtc tca agg gga gta gag ccg ctg gag gcc gcg      222
Gly Ser Gln Asp Glu Val Ser Arg Gly Val Glu Pro Leu Glu Ala Ala
```

-continued

```
 10                  15                  20                  25

cgg gcc cag cct gct aag gac agg agg gcc aag gga acc ccg aag tcc      270
Arg Ala Gln Pro Ala Lys Asp Arg Arg Ala Lys Gly Thr Pro Lys Ser
                 30                  35                  40

tcg aag ccc ggg aaa aaa cac cgg tat ctg aga cta ctt cca gag gcc      318
Ser Lys Pro Gly Lys Lys His Arg Tyr Leu Arg Leu Leu Pro Glu Ala
             45                  50                  55

ttg ata agg ttc ggc ggt ttc cga aaa agg aaa aaa gcc aag tcc tca      366
Leu Ile Arg Phe Gly Gly Phe Arg Lys Arg Lys Lys Ala Lys Ser Ser
         60                  65                  70

gtt tcc aag aag ccg gga gaa gtg gat gac agt ttg gag cag ccc tgt      414
Val Ser Lys Lys Pro Gly Glu Val Asp Asp Ser Leu Glu Gln Pro Cys
     75                  80                  85

ggt ttg ggc tgc tta gtc agc acc tgc tgt ggg tgt tgc aat aac att      462
Gly Leu Gly Cys Leu Val Ser Thr Cys Cys Gly Cys Cys Asn Asn Ile
 90                  95                 100                 105

cgc tgc ttc atg att ttc tac tgc atc ctg ctc ata tgt caa ggt gtg      510
Arg Cys Phe Met Ile Phe Tyr Cys Ile Leu Leu Ile Cys Gln Gly Val
                110                 115                 120

gtg ttt ggt ctt ata gat gtc agc att ggc gat ttt cag aag gaa tat      558
Val Phe Gly Leu Ile Asp Val Ser Ile Gly Asp Phe Gln Lys Glu Tyr
            125                 130                 135

caa ctg aaa acc att gag aag ttg gca ttg gaa aag agt tac gat att      606
Gln Leu Lys Thr Ile Glu Lys Leu Ala Leu Glu Lys Ser Tyr Asp Ile
        140                 145                 150

tca tct ggc ctg gta gca ata ttt ata gca ttc tat gga gac aga aaa      654
Ser Ser Gly Leu Val Ala Ile Phe Ile Ala Phe Tyr Gly Asp Arg Lys
    155                 160                 165

aaa gta ata tgg ttt gta gct tcc tcc ttt tta ata gga ctt gga tca      702
Lys Val Ile Trp Phe Val Ala Ser Ser Phe Leu Ile Gly Leu Gly Ser
170                 175                 180                 185

ctt tta tgt gct ttt cca tcc att aat gaa gaa aat aaa caa agt aag      750
Leu Leu Cys Ala Phe Pro Ser Ile Asn Glu Glu Asn Lys Gln Ser Lys
                190                 195                 200

gta gga att gaa gat att tgc gaa gaa ata aag gtt gtc agt ggt tgc      798
Val Gly Ile Glu Asp Ile Cys Glu Glu Ile Lys Val Val Ser Gly Cys
            205                 210                 215

cag agc agt ggt ata tca ttc caa tca aaa tac ctg tct ttc ttc atc      846
Gln Ser Ser Gly Ile Ser Phe Gln Ser Lys Tyr Leu Ser Phe Phe Ile
        220                 225                 230

ctt ggg cag act gtg cag gga ata gca gga atg cct ctt tat atc ctt      894
Leu Gly Gln Thr Val Gln Gly Ile Ala Gly Met Pro Leu Tyr Ile Leu
    235                 240                 245

gga ata acc ttt att gat gag aat gtt gct aca cac tca gct ggt atc      942
Gly Ile Thr Phe Ile Asp Glu Asn Val Ala Thr His Ser Ala Gly Ile
250                 255                 260                 265

tat tta ggt att gca gaa tgt aca tca atg att gga tat gct ctg ggt      990
Tyr Leu Gly Ile Ala Glu Cys Thr Ser Met Ile Gly Tyr Ala Leu Gly
                270                 275                 280

tat gtg cta gga gca cca cta gtt aaa gtc cct gag aat act act tct     1038
Tyr Val Leu Gly Ala Pro Leu Val Lys Val Pro Glu Asn Thr Thr Ser
            285                 290                 295

gca aca aac act aca gtc aat aat ggt agt cca gaa tgg cta tgg act     1086
Ala Thr Asn Thr Thr Val Asn Asn Gly Ser Pro Glu Trp Leu Trp Thr
        300                 305                 310

tgg tgg att aat ttt ctt ttt gcc gct gtc gtt gca tgg tgt aca tta     1134
Trp Trp Ile Asn Phe Leu Phe Ala Ala Val Val Ala Trp Cys Thr Leu
```

-continued

```
    315                 320                 325

ata cca ttg tca tgc ttt cca aac aat atg cca ggt tca aca cgg ata     1182
Ile Pro Leu Ser Cys Phe Pro Asn Asn Met Pro Gly Ser Thr Arg Ile
330                 335                 340                 345

aaa gct agg aaa cgt aaa cag ctt cat ttt ttt gac agc aga ctt aaa     1230
Lys Ala Arg Lys Arg Lys Gln Leu His Phe Phe Asp Ser Arg Leu Lys
                350                 355                 360

gat ctg aaa ctt gga act aat atc aag gat tta tgt gct gct ctt tgg     1278
Asp Leu Lys Leu Gly Thr Asn Ile Lys Asp Leu Cys Ala Ala Leu Trp
            365                 370                 375

att ctg atg aag aat cca gtg ctc ata tgc cta gct ctg tca aaa gct     1326
Ile Leu Met Lys Asn Pro Val Leu Ile Cys Leu Ala Leu Ser Lys Ala
        380                 385                 390

aca gaa tat tta gtt att att gga gct tct gaa ttt ttg cct ata tat     1374
Thr Glu Tyr Leu Val Ile Ile Gly Ala Ser Glu Phe Leu Pro Ile Tyr
    395                 400                 405

tta gaa aat cag ttt ata tta aca ccc act gtg gca act aca ctt gca     1422
Leu Glu Asn Gln Phe Ile Leu Thr Pro Thr Val Ala Thr Thr Leu Ala
410                 415                 420                 425

gga ctt gtt tta att cca gga ggt gca ctt ggc cag ctt ctg gga ggt     1470
Gly Leu Val Leu Ile Pro Gly Gly Ala Leu Gly Gln Leu Leu Gly Gly
                430                 435                 440

gtc att gtt tcc aca tta gaa atg tct tgt aaa gcc ctt atg aga ttt     1518
Val Ile Val Ser Thr Leu Glu Met Ser Cys Lys Ala Leu Met Arg Phe
            445                 450                 455

ata atg gtt aca tct gtg ata tca ctt ata ctg ctt gtg ttt att att     1566
Ile Met Val Thr Ser Val Ile Ser Leu Ile Leu Leu Val Phe Ile Ile
        460                 465                 470

ttt gta cgc tgt aat cca gtg caa ttt gct ggg atc aat gaa gat tat     1614
Phe Val Arg Cys Asn Pro Val Gln Phe Ala Gly Ile Asn Glu Asp Tyr
    475                 480                 485

gat gga aca ggg aag ttg gga aac ctc acg gct cct tgc aat gaa aaa     1662
Asp Gly Thr Gly Lys Leu Gly Asn Leu Thr Ala Pro Cys Asn Glu Lys
490                 495                 500                 505

tgt aga tgc tca tct tca att tat tct tct ata tgt gga aga gat gat     1710
Cys Arg Cys Ser Ser Ser Ile Tyr Ser Ser Ile Cys Gly Arg Asp Asp
                510                 515                 520

att gaa tat ttt tct ccc tgc ttt gca ggg tgt aca tat tct aaa gca     1758
Ile Glu Tyr Phe Ser Pro Cys Phe Ala Gly Cys Thr Tyr Ser Lys Ala
            525                 530                 535

caa aac caa aaa aag atg tac tac aat tgt tct tgc att aaa gaa gga     1806
Gln Asn Gln Lys Lys Met Tyr Tyr Asn Cys Ser Cys Ile Lys Glu Gly
        540                 545                 550

tta ata act gca gat gca gaa ggt gat ttt att gat gcc aga ccc ggg     1854
Leu Ile Thr Ala Asp Ala Glu Gly Asp Phe Ile Asp Ala Arg Pro Gly
    555                 560                 565

aaa tgt gat gca aag tgc tat aag tta cct ttg ttc att gct ttt atc     1902
Lys Cys Asp Ala Lys Cys Tyr Lys Leu Pro Leu Phe Ile Ala Phe Ile
570                 575                 580                 585

ttt tct aca ctt ata ttt tct ggt ttt tct ggt gta cca atc gtc ttg     1950
Phe Ser Thr Leu Ile Phe Ser Gly Phe Ser Gly Val Pro Ile Val Leu
                590                 595                 600

gcc atg acg cgg gtt gta cct gac aaa ctg cgt tct ctg gcc ttg ggt     1998
Ala Met Thr Arg Val Val Pro Asp Lys Leu Arg Ser Leu Ala Leu Gly
            605                 610                 615

gta agc tat gtg att ttg aga ata ttt ggg act att cct gga cca tca     2046
Val Ser Tyr Val Ile Leu Arg Ile Phe Gly Thr Ile Pro Gly Pro Ser
```

-continued

```
        620                 625                 630

atc ttt aaa atg tca gga gaa act tct tgt att tta cgg gat gtt aat     2094
Ile Phe Lys Met Ser Gly Glu Thr Ser Cys Ile Leu Arg Asp Val Asn
    635                 640                 645

aaa tgt gga cac aga gga cgt tgt tgg ata tat aac aag aca aaa atg     2142
Lys Cys Gly His Arg Gly Arg Cys Trp Ile Tyr Asn Lys Thr Lys Met
650                 655                 660                 665

gct ttc tta ttg gta gga ata tgt ttt ctt tgc aaa cta tgc act atc     2190
Ala Phe Leu Leu Val Gly Ile Cys Phe Leu Cys Lys Leu Cys Thr Ile
                670                 675                 680

atc ttc act act att gca ttt ttc ata tac aaa cgt cgt cta aat gag     2238
Ile Phe Thr Thr Ile Ala Phe Phe Ile Tyr Lys Arg Arg Leu Asn Glu
            685                 690                 695

aac act gac ttc cca gat gta act gtg aag aat cca aaa gtt aag aaa     2286
Asn Thr Asp Phe Pro Asp Val Thr Val Lys Asn Pro Lys Val Lys Lys
        700                 705                 710

aaa gaa gaa act gac ttg taa ctggatcatc attgtattct ccaagatttg        2337
Lys Glu Glu Thr Asp Leu  *
    715

tttctgtgcc caactttcag aagaggaaaa tcacacatta tgtttacata agtagcaaaa   2397

atatatttat ggtgatctgc attttcataa taaagtgtcc tattgtgaaa caaaaaaaaa   2457

aaaaaaaaaa aaagggcggc cgc                                           2480


<210> SEQ ID NO 39
<211> LENGTH: 719
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 39

Met Phe Val Gly Val Ala Arg His Ser Gly Ser Gln Asp Glu Val Ser
 1               5                  10                  15

Arg Gly Val Glu Pro Leu Glu Ala Ala Arg Ala Gln Pro Ala Lys Asp
            20                  25                  30

Arg Arg Ala Lys Gly Thr Pro Lys Ser Ser Lys Pro Gly Lys Lys His
        35                  40                  45

Arg Tyr Leu Arg Leu Leu Pro Glu Ala Leu Ile Arg Phe Gly Gly Phe
    50                  55                  60

Arg Lys Arg Lys Lys Ala Lys Ser Ser Val Ser Lys Lys Pro Gly Glu
65                  70                  75                  80

Val Asp Asp Ser Leu Glu Gln Pro Cys Gly Leu Gly Cys Leu Val Ser
                85                  90                  95

Thr Cys Cys Gly Cys Cys Asn Asn Ile Arg Cys Phe Met Ile Phe Tyr
            100                 105                 110

Cys Ile Leu Leu Ile Cys Gln Gly Val Val Phe Gly Leu Ile Asp Val
        115                 120                 125

Ser Ile Gly Asp Phe Gln Lys Glu Tyr Gln Leu Lys Thr Ile Glu Lys
    130                 135                 140

Leu Ala Leu Glu Lys Ser Tyr Asp Ile Ser Ser Gly Leu Val Ala Ile
145                 150                 155                 160

Phe Ile Ala Phe Tyr Gly Asp Arg Lys Lys Val Ile Trp Phe Val Ala
                165                 170                 175

Ser Ser Phe Leu Ile Gly Leu Gly Ser Leu Leu Cys Ala Phe Pro Ser
            180                 185                 190
```

-continued

```
Ile Asn Glu Glu Asn Lys Gln Ser Lys Val Gly Ile Glu Asp Ile Cys
        195                 200                 205

Glu Glu Ile Lys Val Val Ser Gly Cys Gln Ser Ser Gly Ile Ser Phe
    210                 215                 220

Gln Ser Lys Tyr Leu Ser Phe Phe Ile Leu Gly Gln Thr Val Gln Gly
225                 230                 235                 240

Ile Ala Gly Met Pro Leu Tyr Ile Leu Gly Ile Thr Phe Ile Asp Glu
                245                 250                 255

Asn Val Ala Thr His Ser Ala Gly Ile Tyr Leu Gly Ile Ala Glu Cys
            260                 265                 270

Thr Ser Met Ile Gly Tyr Ala Leu Gly Tyr Val Leu Gly Ala Pro Leu
        275                 280                 285

Val Lys Val Pro Glu Asn Thr Thr Ser Ala Thr Asn Thr Thr Val Asn
    290                 295                 300

Asn Gly Ser Pro Glu Trp Leu Trp Thr Trp Trp Ile Asn Phe Leu Phe
305                 310                 315                 320

Ala Ala Val Val Ala Trp Cys Thr Leu Ile Pro Leu Ser Cys Phe Pro
                325                 330                 335

Asn Asn Met Pro Gly Ser Thr Arg Ile Lys Ala Arg Lys Arg Lys Gln
            340                 345                 350

Leu His Phe Phe Asp Ser Arg Leu Lys Asp Leu Lys Leu Gly Thr Asn
        355                 360                 365

Ile Lys Asp Leu Cys Ala Ala Leu Trp Ile Leu Met Lys Asn Pro Val
    370                 375                 380

Leu Ile Cys Leu Ala Leu Ser Lys Ala Thr Glu Tyr Leu Val Ile Ile
385                 390                 395                 400

Gly Ala Ser Glu Phe Leu Pro Ile Tyr Leu Glu Asn Gln Phe Ile Leu
                405                 410                 415

Thr Pro Thr Val Ala Thr Thr Leu Ala Gly Leu Val Leu Ile Pro Gly
            420                 425                 430

Gly Ala Leu Gly Gln Leu Leu Gly Gly Val Ile Val Ser Thr Leu Glu
        435                 440                 445

Met Ser Cys Lys Ala Leu Met Arg Phe Ile Met Val Thr Ser Val Ile
    450                 455                 460

Ser Leu Ile Leu Leu Val Phe Ile Ile Phe Val Arg Cys Asn Pro Val
465                 470                 475                 480

Gln Phe Ala Gly Ile Asn Glu Asp Tyr Asp Gly Thr Gly Lys Leu Gly
                485                 490                 495

Asn Leu Thr Ala Pro Cys Asn Glu Lys Cys Arg Cys Ser Ser Ser Ile
            500                 505                 510

Tyr Ser Ser Ile Cys Gly Arg Asp Asp Ile Glu Tyr Phe Ser Pro Cys
        515                 520                 525

Phe Ala Gly Cys Thr Tyr Ser Lys Ala Gln Asn Gln Lys Lys Met Tyr
    530                 535                 540

Tyr Asn Cys Ser Cys Ile Lys Glu Gly Leu Ile Thr Ala Asp Ala Glu
545                 550                 555                 560

Gly Asp Phe Ile Asp Ala Arg Pro Gly Lys Cys Asp Ala Lys Cys Tyr
                565                 570                 575

Lys Leu Pro Leu Phe Ile Ala Phe Ile Phe Ser Thr Leu Ile Phe Ser
            580                 585                 590

Gly Phe Ser Gly Val Pro Ile Val Leu Ala Met Thr Arg Val Val Pro
```

-continued

```
        595                 600                 605

Asp Lys Leu Arg Ser Leu Ala Leu Gly Val Ser Tyr Val Ile Leu Arg
    610                 615                 620

Ile Phe Gly Thr Ile Pro Gly Pro Ser Ile Phe Lys Met Ser Gly Glu
625                 630                 635                 640

Thr Ser Cys Ile Leu Arg Asp Val Asn Lys Cys Gly His Arg Gly Arg
                645                 650                 655

Cys Trp Ile Tyr Asn Lys Thr Lys Met Ala Phe Leu Leu Val Gly Ile
            660                 665                 670

Cys Phe Leu Cys Lys Leu Cys Thr Ile Ile Phe Thr Thr Ile Ala Phe
        675                 680                 685

Phe Ile Tyr Lys Arg Arg Leu Asn Glu Asn Thr Asp Phe Pro Asp Val
    690                 695                 700

Thr Val Lys Asn Pro Lys Val Lys Lys Lys Glu Glu Thr Asp Leu
705                 710                 715


<210> SEQ ID NO 40
<211> LENGTH: 2160
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(2160)

<400> SEQUENCE: 40

atg ttc gta ggc gtc gcc cgg cac tct ggg agc cag gat gaa gtc tca       48
Met Phe Val Gly Val Ala Arg His Ser Gly Ser Gln Asp Glu Val Ser
 1               5                  10                  15

agg gga gta gag ccg ctg gag gcc gcg cgg gcc cag cct gct aag gac       96
Arg Gly Val Glu Pro Leu Glu Ala Ala Arg Ala Gln Pro Ala Lys Asp
            20                  25                  30

agg agg gcc aag gga acc ccg aag tcc tcg aag ccc ggg aaa aaa cac      144
Arg Arg Ala Lys Gly Thr Pro Lys Ser Ser Lys Pro Gly Lys Lys His
        35                  40                  45

cgg tat ctg aga cta ctt cca gag gcc ttg ata agg ttc ggc ggt ttc      192
Arg Tyr Leu Arg Leu Leu Pro Glu Ala Leu Ile Arg Phe Gly Gly Phe
    50                  55                  60

cga aaa agg aaa aaa gcc aag tcc tca gtt tcc aag aag ccg gga gaa      240
Arg Lys Arg Lys Lys Ala Lys Ser Ser Val Ser Lys Lys Pro Gly Glu
 65                  70                  75                  80

gtg gat gac agt ttg gag cag ccc tgt ggt ttg ggc tgc tta gtc agc      288
Val Asp Asp Ser Leu Glu Gln Pro Cys Gly Leu Gly Cys Leu Val Ser
                85                  90                  95

acc tgc tgt ggg tgt tgc aat aac att cgc tgc ttc atg att ttc tac      336
Thr Cys Cys Gly Cys Cys Asn Asn Ile Arg Cys Phe Met Ile Phe Tyr
            100                 105                 110

tgc atc ctg ctc ata tgt caa ggt gtg gtg ttt ggt ctt ata gat gtc      384
Cys Ile Leu Leu Ile Cys Gln Gly Val Val Phe Gly Leu Ile Asp Val
        115                 120                 125

agc att ggc gat ttt cag aag gaa tat caa ctg aaa acc att gag aag      432
Ser Ile Gly Asp Phe Gln Lys Glu Tyr Gln Leu Lys Thr Ile Glu Lys
    130                 135                 140

ttg gca ttg gaa aag agt tac gat att tca tct ggc ctg gta gca ata      480
Leu Ala Leu Glu Lys Ser Tyr Asp Ile Ser Ser Gly Leu Val Ala Ile
145                 150                 155                 160

ttt ata gca ttc tat gga gac aga aaa aaa gta ata tgg ttt gta gct      528
Phe Ile Ala Phe Tyr Gly Asp Arg Lys Lys Val Ile Trp Phe Val Ala
```

-continued

```
                165                 170                 175

tcc tcc ttt tta ata gga ctt gga tca ctt tta tgt gct ttt cca tcc      576
Ser Ser Phe Leu Ile Gly Leu Gly Ser Leu Leu Cys Ala Phe Pro Ser
            180                 185                 190

att aat gaa gaa aat aaa caa agt aag gta gga att gaa gat att tgc      624
Ile Asn Glu Glu Asn Lys Gln Ser Lys Val Gly Ile Glu Asp Ile Cys
        195                 200                 205

gaa gaa ata aag gtt gtc agt ggt tgc cag agc agt ggt ata tca ttc      672
Glu Glu Ile Lys Val Val Ser Gly Cys Gln Ser Ser Gly Ile Ser Phe
    210                 215                 220

caa tca aaa tac ctg tct ttc ttc atc ctt ggg cag act gtg cag gga      720
Gln Ser Lys Tyr Leu Ser Phe Phe Ile Leu Gly Gln Thr Val Gln Gly
225                 230                 235                 240

ata gca gga atg cct ctt tat atc ctt gga ata acc ttt att gat gag      768
Ile Ala Gly Met Pro Leu Tyr Ile Leu Gly Ile Thr Phe Ile Asp Glu
                245                 250                 255

aat gtt gct aca cac tca gct ggt atc tat tta ggt att gca gaa tgt      816
Asn Val Ala Thr His Ser Ala Gly Ile Tyr Leu Gly Ile Ala Glu Cys
            260                 265                 270

aca tca atg att gga tat gct ctg ggt tat gtg cta gga gca cca cta      864
Thr Ser Met Ile Gly Tyr Ala Leu Gly Tyr Val Leu Gly Ala Pro Leu
        275                 280                 285

gtt aaa gtc cct gag aat act act tct gca aca aac act aca gtc aat      912
Val Lys Val Pro Glu Asn Thr Thr Ser Ala Thr Asn Thr Thr Val Asn
    290                 295                 300

aat ggt agt cca gaa tgg cta tgg act tgg tgg att aat ttt ctt ttt      960
Asn Gly Ser Pro Glu Trp Leu Trp Thr Trp Trp Ile Asn Phe Leu Phe
305                 310                 315                 320

gcc gct gtc gtt gca tgg tgt aca tta ata cca ttg tca tgc ttt cca     1008
Ala Ala Val Val Ala Trp Cys Thr Leu Ile Pro Leu Ser Cys Phe Pro
                325                 330                 335

aac aat atg cca ggt tca aca cgg ata aaa gct agg aaa cgt aaa cag     1056
Asn Asn Met Pro Gly Ser Thr Arg Ile Lys Ala Arg Lys Arg Lys Gln
            340                 345                 350

ctt cat ttt ttt gac agc aga ctt aaa gat ctg aaa ctt gga act aat     1104
Leu His Phe Phe Asp Ser Arg Leu Lys Asp Leu Lys Leu Gly Thr Asn
        355                 360                 365

atc aag gat tta tgt gct gct ctt tgg att ctg atg aag aat cca gtg     1152
Ile Lys Asp Leu Cys Ala Ala Leu Trp Ile Leu Met Lys Asn Pro Val
    370                 375                 380

ctc ata tgc cta gct ctg tca aaa gct aca gaa tat tta gtt att att     1200
Leu Ile Cys Leu Ala Leu Ser Lys Ala Thr Glu Tyr Leu Val Ile Ile
385                 390                 395                 400

gga gct tct gaa ttt ttg cct ata tat tta gaa aat cag ttt ata tta     1248
Gly Ala Ser Glu Phe Leu Pro Ile Tyr Leu Glu Asn Gln Phe Ile Leu
                405                 410                 415

aca ccc act gtg gca act aca ctt gca gga ctt gtt tta att cca gga     1296
Thr Pro Thr Val Ala Thr Thr Leu Ala Gly Leu Val Leu Ile Pro Gly
            420                 425                 430

ggt gca ctt ggc cag ctt ctg gga ggt gtc att gtt tcc aca tta gaa     1344
Gly Ala Leu Gly Gln Leu Leu Gly Gly Val Ile Val Ser Thr Leu Glu
        435                 440                 445

atg tct tgt aaa gcc ctt atg aga ttt ata atg gtt aca tct gtg ata     1392
Met Ser Cys Lys Ala Leu Met Arg Phe Ile Met Val Thr Ser Val Ile
    450                 455                 460

tca ctt ata ctg ctt gtg ttt att att ttt gta cgc tgt aat cca gtg     1440
Ser Leu Ile Leu Leu Val Phe Ile Ile Phe Val Arg Cys Asn Pro Val
```

-continued

```
465                 470                 475                 480

caa ttt gct ggg atc aat gaa gat tat gat gga aca ggg aag ttg gga     1488
Gln Phe Ala Gly Ile Asn Glu Asp Tyr Asp Gly Thr Gly Lys Leu Gly
                485                 490                 495

aac ctc acg gct cct tgc aat gaa aaa tgt aga tgc tca tct tca att     1536
Asn Leu Thr Ala Pro Cys Asn Glu Lys Cys Arg Cys Ser Ser Ser Ile
            500                 505                 510

tat tct tct ata tgt gga aga gat gat att gaa tat ttt tct ccc tgc     1584
Tyr Ser Ser Ile Cys Gly Arg Asp Asp Ile Glu Tyr Phe Ser Pro Cys
        515                 520                 525

ttt gca ggg tgt aca tat tct aaa gca caa aac caa aaa aag atg tac     1632
Phe Ala Gly Cys Thr Tyr Ser Lys Ala Gln Asn Gln Lys Lys Met Tyr
    530                 535                 540

tac aat tgt tct tgc att aaa gaa gga tta ata act gca gat gca gaa     1680
Tyr Asn Cys Ser Cys Ile Lys Glu Gly Leu Ile Thr Ala Asp Ala Glu
545                 550                 555                 560

ggt gat ttt att gat gcc aga ccc ggg aaa tgt gat gca aag tgc tat     1728
Gly Asp Phe Ile Asp Ala Arg Pro Gly Lys Cys Asp Ala Lys Cys Tyr
                565                 570                 575

aag tta cct ttg ttc att gct ttt atc ttt tct aca ctt ata ttt tct     1776
Lys Leu Pro Leu Phe Ile Ala Phe Ile Phe Ser Thr Leu Ile Phe Ser
            580                 585                 590

ggt ttt tct ggt gta cca atc gtc ttg gcc atg acg cgg gtt gta cct     1824
Gly Phe Ser Gly Val Pro Ile Val Leu Ala Met Thr Arg Val Val Pro
        595                 600                 605

gac aaa ctg cgt tct ctg gcc ttg ggt gta agc tat gtg att ttg aga     1872
Asp Lys Leu Arg Ser Leu Ala Leu Gly Val Ser Tyr Val Ile Leu Arg
    610                 615                 620

ata ttt ggg act att cct gga cca tca atc ttt aaa atg tca gga gaa     1920
Ile Phe Gly Thr Ile Pro Gly Pro Ser Ile Phe Lys Met Ser Gly Glu
625                 630                 635                 640

act tct tgt att tta cgg gat gtt aat aaa tgt gga cac aga gga cgt     1968
Thr Ser Cys Ile Leu Arg Asp Val Asn Lys Cys Gly His Arg Gly Arg
                645                 650                 655

tgt tgg ata tat aac aag aca aaa atg gct ttc tta ttg gta gga ata     2016
Cys Trp Ile Tyr Asn Lys Thr Lys Met Ala Phe Leu Leu Val Gly Ile
            660                 665                 670

tgt ttt ctt tgc aaa cta tgc act atc atc ttc act act att gca ttt     2064
Cys Phe Leu Cys Lys Leu Cys Thr Ile Ile Phe Thr Thr Ile Ala Phe
        675                 680                 685

ttc ata tac aaa cgt cgt cta aat gag aac act gac ttc cca gat gta     2112
Phe Ile Tyr Lys Arg Arg Leu Asn Glu Asn Thr Asp Phe Pro Asp Val
    690                 695                 700

act gtg aag aat cca aaa gtt aag aaa aaa gaa gaa act gac ttg taa     2160
Thr Val Lys Asn Pro Lys Val Lys Lys Lys Glu Glu Thr Asp Leu  *
705                 710                 715


<210> SEQ ID NO 41
<211> LENGTH: 61
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 41

Cys Ser Arg Tyr Pro Asn Pro Thr Ser Lys Asp Gly Lys Leu Val Ala
 1               5                  10                  15

Cys Pro Arg Glu Tyr Asp Pro Val Cys Gly Ser Asp Gly Val Thr Tyr
```

-continued

```
            20                  25                  30

Ser Asn Glu Cys Glu Leu Lys Lys Ala Ala Cys Ala Glu Asn Val Glu
        35                  40                  45

Gln Gly Thr Asn Ile Glu Lys Lys His Asp Gly Pro Cys
    50                  55                  60


<210> SEQ ID NO 42
<211> LENGTH: 58
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 42

Asp Cys Ser Glu Tyr Arg Ser Pro Thr Ser Gly Gly Leu Leu Ala Cys
 1               5                  10                  15

Pro Arg Glu Tyr Asp Pro Val Cys Gly Ser Asp Gly Val Thr Tyr Ser
            20                  25                  30

Asn Glu Cys Glu Leu Cys Ala Ala Ala Cys Glu Ala Glu Gln Gly Lys
        35                  40                  45

Ser Ile Glu Val Lys His Asp Gly Pro Cys
    50                  55


<210> SEQ ID NO 43
<211> LENGTH: 192
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 43

Cys Asn Glu Gln Cys Ser Cys Glu Thr Ser Thr Trp His Pro Val Cys
 1               5                  10                  15

Gly Asp Asp Asn Gly Leu Ala Tyr Tyr Ser Pro Cys His Ala Gly Cys
            20                  25                  30

Ser Glu Thr Asn Gln Ser Ser Gly Thr Gly Thr Asn Met Val Phe Thr
        35                  40                  45

Asn Cys Ser Cys Val Gln Thr Ser Gly Asn Ser Ser Ala Val Val Lys
    50                  55                  60

Lys Cys Cys Lys Asn Pro Glu Cys Gln Asn Lys Leu Gln Tyr Phe Leu
65                  70                  75                  80

Ile Leu Met Ile Phe Gly Ser Phe Ile Tyr Ser Leu Ala Ala Val Pro
                85                  90                  95

Gly Tyr Met Val Ile Leu Arg Cys Val Asn Pro Glu Glu Lys Ser Leu
            100                 105                 110

Ala Leu Gly Leu His Trp Phe Cys Val Arg Ile Phe Gly Thr Ile Pro
        115                 120                 125

Ala Pro Ile Ile Phe Gly Leu Leu Ile Asp Met Ser Cys Leu His Trp
    130                 135                 140

Asn Lys Gln Cys Cys Gly Glu Arg Gly Ser Cys Arg Met Tyr Asp Asn
145                 150                 155                 160

Asp Ser Leu Arg Asn Met Tyr His Gly Leu Thr Val Ala Leu Arg Val
                165                 170                 175

Ile Ser Leu Ile Pro Ala Phe Phe Val Trp Trp Leu Met Lys Lys Asn
            180                 185                 190
```

-continued

<210> SEQ ID NO 44
<211> LENGTH: 712
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 44

Met Asp Thr Ser Ser Lys Glu Asn Ile Gln Leu Phe Cys Lys Thr Ser
1 5 10 15

Val Gln Pro Val Gly Arg Pro Ser Phe Lys Thr Glu Tyr Pro Ser Ser
20 25 30

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

Leu Ser Phe Val Tyr Phe Ala Lys Ala Leu Ala Glu Gly Tyr Leu Lys
50 55 60

Ser Thr Ile Thr Gln Ile Glu Arg Arg Phe Asp Ile Pro Ser Ser Leu
65 70 75 80

Val Gly Val Ile Asp Gly Ser Phe Glu Ile Gly Asn Leu Leu Val Ile
85 90 95

Thr Phe Val Ser Tyr Phe Gly Ala Lys Leu His Arg Pro Lys Ile Ile
100 105 110

Gly Ala Gly Cys Val Ile Met Gly Val Gly Thr Leu Leu Ile Ala Met
115 120 125

Pro Gln Phe Phe Met Glu Gln Tyr Lys Tyr Glu Arg Tyr Ser Pro Ser
130 135 140

Ser Asn Ser Thr Leu Ser Ile Ser Pro Cys Leu Leu Glu Ser Ser Ser
145 150 155 160

Gln Leu Pro Val Ser Val Met Glu Lys Ser Lys Ser Lys Ile Ser Asn
165 170 175

Glu Cys Glu Val Asp Thr Ser Ser Ser Met Trp Ile Tyr Val Phe Leu
180 185 190

Gly Asn Leu Leu Arg Gly Ile Gly Glu Thr Pro Ile Gln Pro Leu Gly
195 200 205

Ile Ala Tyr Leu Asp Asp Phe Ala Ser Glu Asp Asn Ala Ala Phe Tyr
210 215 220

Ile Gly Cys Val Gln Thr Val Ala Ile Ile Gly Pro Ile Phe Gly Phe
225 230 235 240

Leu Leu Gly Ser Leu Cys Ala Lys Leu Tyr Val Asp Ile Gly Phe Val
245 250 255

Asn Leu Asp His Ile Thr Ile Thr Pro Lys Asp Pro Gln Trp Val Gly
260 265 270

Ala Trp Trp Leu Gly Tyr Leu Ile Ala Gly Ile Ile Ser Leu Leu Ala
275 280 285

Ala Val Pro Phe Trp Tyr Leu Pro Lys Ser Leu Pro Arg Ser Gln Ser
290 295 300

Arg Glu Asp Ser Asn Ser Ser Ser Glu Lys Ser Lys Phe Ile Ile Asp
305 310 315 320

Asp His Thr Asp Tyr Gln Thr Pro Gln Gly Glu Asn Ala Lys Ile Met
325 330 335

Glu Met Ala Arg Asp Phe Leu Pro Ser Leu Lys Asn Leu Phe Gly Asn
340 345 350

Pro Val Tyr Phe Leu Tyr Leu Cys Thr Ser Thr Val Gln Phe Asn Ser
355 360 365

-continued

```
Leu Phe Gly Met Val Thr Tyr Lys Pro Lys Tyr Ile Glu Gln Gln Tyr
    370                 375                 380

Gly Gln Ser Ser Ser Arg Ala Asn Phe Val Ile Gly Leu Ile Asn Ile
385                 390                 395                 400

Pro Ala Val Ala Leu Gly Ile Phe Ser Gly Gly Ile Val Met Lys Lys
                405                 410                 415

Phe Arg Ile Ser Val Cys Gly Ala Ala Lys Leu Tyr Leu Gly Ser Ser
            420                 425                 430

Val Phe Gly Tyr Leu Leu Phe Leu Ser Leu Phe Ala Leu Gly Cys Glu
        435                 440                 445

Asn Ser Asp Val Ala Gly Leu Thr Val Ser Tyr Gln Gly Thr Lys Pro
    450                 455                 460

Val Ser Tyr His Glu Arg Ala Leu Phe Ser Asp Cys Asn Ser Arg Cys
465                 470                 475                 480

Lys Cys Ser Glu Thr Lys Trp Glu Pro Met Cys Gly Glu Asn Gly Ile
                485                 490                 495

Thr Tyr Val Ser Ala Cys Leu Ala Gly Cys Gln Thr Ser Asn Arg Ser
            500                 505                 510

Gly Lys Asn Ile Ile Phe Tyr Asn Cys Thr Cys Val Gly Ile Ala Ala
        515                 520                 525

Ser Lys Ser Gly Asn Ser Ser Gly Ile Val Gly Arg Cys Gln Lys Asp
    530                 535                 540

Asn Gly Cys Pro Gln Met Phe Leu Tyr Phe Leu Val Ile Ser Val Ile
545                 550                 555                 560

Thr Ser Tyr Thr Leu Ser Leu Gly Gly Ile Pro Gly Tyr Ile Leu Leu
                565                 570                 575

Leu Arg Cys Ile Lys Pro Gln Leu Lys Ser Phe Ala Leu Gly Ile Tyr
            580                 585                 590

Thr Leu Ala Ile Arg Val Leu Ala Gly Ile Pro Ala Pro Val Tyr Phe
        595                 600                 605

Gly Val Leu Ile Asp Thr Ser Cys Leu Lys Trp Gly Phe Lys Arg Cys
    610                 615                 620

Gly Ser Arg Gly Ser Cys Arg Leu Tyr Asp Ser Asn Val Phe Arg His
625                 630                 635                 640

Ile Tyr Leu Gly Leu Thr Val Ile Leu Gly Thr Val Ser Ile Leu Leu
                645                 650                 655

Ser Ile Ala Val Leu Phe Ile Leu Lys Lys Asn Tyr Val Ser Lys His
            660                 665                 670

Arg Ser Phe Ile Thr Lys Arg Glu Arg Thr Met Val Ser Thr Arg Phe
        675                 680                 685

Gln Lys Glu Asn Tyr Thr Thr Ser Asp His Leu Leu Gln Pro Asn Tyr
    690                 695                 700

Trp Pro Gly Lys Glu Thr Gln Leu
705                 710
```

<210> SEQ ID NO 45
<211> LENGTH: 487
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 45

-continued

```
Val Ala Leu Val Ala Ala Leu Gly Gly Gly Phe Leu Phe Gly Tyr Asp
 1               5                  10                  15

Thr Gly Val Ile Gly Gly Phe Leu Ala Leu Ile Asp Phe Leu Phe Arg
            20                  25                  30

Phe Gly Leu Leu Thr Ser Ser Gly Ala Leu Ala Glu Leu Gly Tyr Ser
        35                  40                  45

Thr Val Leu Thr Gly Leu Val Val Ser Ile Phe Phe Leu Gly Arg Leu
    50                  55                  60

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

Lys Ser Leu Leu Ile Ala Leu Val Leu Phe Val Ile Gly Ala Leu Leu
                85                  90                  95

Ser Gly Ala Ala Pro Gly Tyr Thr Thr Ile Gly Leu Trp Ala Phe Tyr
            100                 105                 110

Leu Leu Ile Val Gly Arg Val Leu Val Gly Leu Gly Val Gly Gly Ala
        115                 120                 125

Ser Val Leu Val Pro Met Tyr Ile Ser Glu Ile Ala Pro Lys Ala Leu
    130                 135                 140

Arg Gly Ala Leu Gly Ser Leu Tyr Gln Leu Ala Ile Thr Ile Gly Ile
145                 150                 155                 160

Leu Val Ala Ala Ile Ile Gly Leu Gly Leu Asn Lys Thr Asn Asn Asp
                165                 170                 175

Ser Ala Leu Asn Ser Trp Gly Trp Arg Ile Pro Leu Gly Leu Gln Leu
            180                 185                 190

Val Pro Ala Leu Leu Leu Leu Ile Gly Leu Leu Phe Leu Pro Glu Ser
        195                 200                 205

Pro Arg Trp Leu Val Glu Lys Gly Lys Leu Glu Glu Ala Arg Glu Val
    210                 215                 220

Leu Ala Lys Leu Arg Gly Val Glu Asp Val Asp Gln Glu Ile Gln Glu
225                 230                 235                 240

Ile Lys Ala Glu Leu Glu Ala Gly Val Glu Glu Glu Lys Ala Gly Lys
                245                 250                 255

Ala Ser Trp Gly Glu Leu Phe Arg Gly Arg Thr Arg Pro Lys Val Arg
            260                 265                 270

Gln Arg Leu Leu Met Gly Val Met Leu Gln Ala Phe Gln Gln Leu Thr
        275                 280                 285

Gly Ile Asn Ala Ile Phe Tyr Tyr Ser Pro Thr Ile Phe Lys Ser Val
    290                 295                 300

Gly Val Ser Asp Ser Arg Ala Ser Leu Leu Val Thr Ile Ile Val Gly
305                 310                 315                 320

Val Val Asn Phe Val Phe Thr Leu Val Ala Leu Ile Phe Leu Val Asp
                325                 330                 335

Arg Phe Gly Arg Arg Pro Leu Leu Leu Leu Gly Ala Ala Gly Met Ala
            340                 345                 350

Ile Cys Phe Leu Ile Leu Gly Ala Ser Ile Gly Val Ala Leu Leu Leu
        355                 360                 365

Leu Asn Lys Pro Lys Asp Pro Leu Ser Lys Ala Ala Gly Ile Val Ala
    370                 375                 380

Ile Val Phe Ile Leu Leu Phe Ile Ala Phe Phe Ala Leu Gly Trp Gly
385                 390                 395                 400
```

-continued

```
Pro Ile Pro Trp Val Ile Leu Ser Glu Leu Phe Pro Thr Lys Val Arg
                405                 410                 415

Ser Lys Ala Leu Ala Leu Ala Thr Ala Ala Asn Trp Leu Ala Asn Phe
            420                 425                 430

Ile Ile Gly Phe Leu Phe Pro Tyr Ile Thr Gly Ala Ile Gly Leu Ala
        435                 440                 445

Leu Gly Gly Tyr Val Phe Leu Val Phe Ala Gly Leu Leu Val Leu Phe
    450                 455                 460

Ile Leu Phe Val Phe Phe Phe Val Pro Glu Thr Lys Gly Arg Thr Leu
465                 470                 475                 480

Glu Glu Ile Glu Glu Leu Phe
                485


<210> SEQ ID NO 46
<211> LENGTH: 44
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 46

Asn Gly Thr Glu Pro Asn Gly Ser Asp Pro Val Thr Glu Pro Cys Ser
 1               5                  10                  15

Asp Gly Trp Val Tyr Asp His Ser Thr Phe Pro Ser Thr Ile Val Thr
            20                  25                  30

Glu Trp Asp Leu Val Cys Asn His Arg Ala Leu Arg
        35                  40


<210> SEQ ID NO 47
<211> LENGTH: 550
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 47

Met Ala Phe Ser Lys Leu Leu Glu Gln Ala Gly Gly Val Gly Leu Phe
 1               5                  10                  15

Gln Thr Leu Gln Val Leu Thr Phe Ile Leu Pro Cys Leu Met Ile Pro
            20                  25                  30

Ser Gln Met Leu Leu Glu Asn Phe Ser Ala Ala Ile Pro Gly His Arg
        35                  40                  45

Cys Trp Thr His Met Leu Asp Asn Gly Ser Ala Val Ser Thr Asn Met
    50                  55                  60

Thr Pro Lys Ala Leu Leu Thr Ile Ser Ile Pro Pro Gly Pro Asn Gln
65                  70                  75                  80

Gly Pro His Gln Cys Arg Arg Phe Arg Gln Pro Gln Trp Gln Leu Leu
                85                  90                  95

Asp Pro Asn Ala Thr Ala Thr Ser Trp Ser Glu Ala Asp Thr Glu Pro
            100                 105                 110

Cys Val Asp Gly Trp Val Tyr Asp Arg Ser Val Phe Thr Ser Thr Ile
        115                 120                 125

Val Ala Lys Trp Asp Leu Val Cys Ser Ser Gln Gly Leu Lys Pro Leu
    130                 135                 140

Ser Gln Ser Ile Phe Met Ser Gly Ile Leu Val Gly Ser Phe Ile Trp
145                 150                 155                 160

Gly Leu Leu Ser Tyr Arg Phe Gly Arg Lys Pro Met Leu Ser Trp Cys
```

-continued

```
                165                 170                 175

Cys Leu Gln Leu Ala Val Ala Gly Thr Ser Thr Ile Phe Ala Pro Thr
            180                 185                 190

Phe Val Ile Tyr Cys Gly Leu Arg Phe Val Ala Ala Phe Gly Met Ala
        195                 200                 205

Gly Ile Phe Leu Ser Ser Leu Thr Leu Met Val Glu Trp Thr Thr Thr
    210                 215                 220

Ser Arg Arg Ala Val Thr Met Thr Val Val Gly Cys Ala Phe Ser Ala
225                 230                 235                 240

Gly Gln Ala Ala Leu Gly Gly Leu Ala Phe Ala Leu Arg Asp Trp Arg
                245                 250                 255

Thr Leu Gln Leu Ala Ala Ser Val Pro Phe Phe Ala Ile Ser Leu Ile
            260                 265                 270

Ser Trp Trp Leu Pro Glu Ser Ala Arg Trp Leu Ile Ile Lys Gly Lys
        275                 280                 285

Pro Asp Gln Ala Leu Gln Glu Leu Arg Lys Val Ala Arg Ile Asn Gly
    290                 295                 300

His Lys Glu Ala Lys Asn Leu Thr Ile Glu Val Leu Met Ser Ser Val
305                 310                 315                 320

Lys Glu Glu Val Ala Ser Ala Lys Glu Pro Arg Ser Val Leu Asp Leu
                325                 330                 335

Phe Cys Val Pro Val Leu Arg Trp Arg Ser Cys Ala Met Leu Val Val
            340                 345                 350

Asn Phe Ser Leu Leu Ile Ser Tyr Tyr Gly Leu Val Phe Asp Leu Gln
        355                 360                 365

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

Asp Phe Leu Gly Arg Ala Thr Thr Ala Leu Leu Leu Ser Phe Leu Gly
385                 390                 395                 400

Arg Arg Thr Ile Gln Ala Gly Ser Gln Ala Met Ala Gly Leu Ala Ile
                405                 410                 415

Leu Ala Asn Met Leu Val Pro Gln Asp Leu Gln Thr Leu Arg Val Val
            420                 425                 430

Phe Ala Val Leu Gly Lys Gly Cys Phe Gly Ile Ser Leu Thr Cys Leu
        435                 440                 445

Thr Ile Tyr Lys Ala Glu Leu Phe Pro Thr Pro Val Arg Met Thr Ala
    450                 455                 460

Asp Gly Ile Leu His Thr Val Gly Arg Leu Gly Ala Met Met Gly Pro
465                 470                 475                 480

Leu Ile Leu Met Ser Arg Gln Ala Leu Pro Leu Leu Pro Pro Leu Leu
                485                 490                 495

Tyr Gly Val Ile Ser Ile Ala Ser Ser Leu Val Val Leu Phe Phe Leu
            500                 505                 510

Pro Glu Thr Gln Gly Leu Pro Leu Pro Asp Thr Ile Gln Asp Leu Glu
        515                 520                 525

Ser Gln Lys Ser Thr Ala Ala Gln Gly Asn Arg Gln Glu Ala Val Thr
    530                 535                 540

Val Glu Ser Thr Ser Leu
545                 550
```

<210> SEQ ID NO 48

-continued

<211> LENGTH: 553
<212> TYPE: PRT
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 48

Met Ala Phe Pro Glu Leu Leu Asp Arg Val Gly Gly Leu Gly Arg Phe
1 5 10 15

Gln Leu Phe Gln Thr Val Ala Leu Val Thr Pro Ile Leu Trp Val Thr
20 25 30

Thr Gln Asn Met Leu Glu Asn Phe Ser Ala Ala Val Pro His His Arg
35 40 45

Cys Trp Val Pro Leu Leu Asp Asn Ser Thr Ser Gln Ala Ser Ile Pro
50 55 60

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

Pro Asp Gln Gln Pro His Gln Cys Leu Arg Phe Arg Gln Pro Gln Trp
85 90 95

Gln Leu Thr Glu Ser Asn Ala Thr Ala Thr Asn Trp Ser Asp Ala Ala
100 105 110

Thr Glu Pro Cys Glu Asp Gly Trp Val Tyr Asp His Ser Thr Phe Arg
115 120 125

Ser Thr Ile Val Thr Thr Trp Asp Leu Val Cys Asn Ser Gln Ala Leu
130 135 140

Arg Pro Met Ala Gln Ser Ile Phe Leu Ala Gly Ile Leu Val Gly Ala
145 150 155 160

Ala Val Cys Gly His Ala Ser Asp Arg Phe Gly Arg Arg Arg Val Leu
165 170 175

Thr Trp Ser Tyr Leu Leu Val Ser Val Ser Gly Thr Ala Ala Ala Phe
180 185 190

Met Pro Thr Phe Pro Leu Tyr Cys Leu Phe Arg Phe Leu Leu Ala Ser
195 200 205

Ala Val Ala Gly Val Met Met Asn Thr Ala Ser Leu Leu Met Glu Trp
210 215 220

Thr Ser Ala Gln Gly Ser Pro Leu Val Met Thr Leu Asn Ala Leu Gly
225 230 235 240

Phe Ser Phe Gly Gln Val Leu Thr Gly Ser Val Ala Tyr Gly Val Arg
245 250 255

Ser Trp Arg Met Leu Gln Leu Ala Val Ser Ala Pro Phe Phe Leu Phe
260 265 270

Phe Val Tyr Ser Trp Trp Leu Pro Glu Ser Ala Arg Trp Leu Ile Thr
275 280 285

Val Gly Lys Leu Asp Gln Gly Leu Gln Glu Leu Gln Arg Val Ala Ala
290 295 300

Val Asn Arg Arg Lys Ala Glu Gly Asp Thr Leu Thr Met Glu Val Leu
305 310 315 320

Arg Ser Ala Met Glu Glu Glu Pro Ser Arg Asp Lys Ala Gly Ala Ser
325 330 335

Leu Gly Thr Leu Leu His Thr Pro Gly Leu Arg His Arg Thr Ile Ile
340 345 350

Ser Met Leu Cys Trp Phe Ala Phe Gly Phe Thr Phe Tyr Gly Leu Ala
355 360 365

Leu Asp Leu Gln Ala Leu Gly Ser Asn Ile Phe Leu Leu Gln Ala Leu

-continued

```
    370                 375                 380

Ile Gly Ile Val Asp Phe Pro Val Lys Thr Gly Ser Leu Leu Leu Ile
385                 390                 395                 400

Ser Arg Leu Gly Arg Arg Leu Cys Gln Val Ser Phe Leu Val Leu Pro
                405                 410                 415

Gly Leu Cys Ile Leu Ser Asn Ile Leu Val Pro His Gly Met Gly Val
            420                 425                 430

Leu Arg Ser Ala Leu Ala Val Leu Gly Leu Gly Cys Leu Gly Gly Ala
        435                 440                 445

Phe Thr Cys Ile Thr Ile Phe Ser Ser Glu Leu Phe Pro Thr Val Ile
    450                 455                 460

Arg Met Thr Ala Val Gly Leu Cys Gln Val Ala Ala Arg Gly Gly Ala
465                 470                 475                 480

Met Leu Gly Pro Leu Val Arg Leu Leu Gly Val Tyr Gly Ser Trp Met
                485                 490                 495

Pro Leu Leu Val Tyr Gly Val Val Pro Val Leu Ser Gly Leu Ala Ala
            500                 505                 510

Leu Leu Leu Pro Glu Thr Lys Asn Leu Pro Leu Pro Asp Thr Ile Gln
        515                 520                 525

Asp Ile Gln Lys Gln Ser Val Lys Lys Val Thr His Asp Thr Pro Asp
    530                 535                 540

Gly Ser Ile Leu Met Ser Thr Arg Leu
545                 550


<210> SEQ ID NO 49
<211> LENGTH: 722
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 49

Met Pro Leu His Gln Leu Gly Asp Lys Pro Leu Thr Phe Pro Ser Pro
 1               5                  10                  15

Asn Ser Ala Met Glu Asn Gly Leu Asp His Thr Pro Pro Ser Arg Arg
            20                  25                  30

Ala Ser Pro Gly Thr Pro Leu Ser Pro Gly Ser Leu Arg Ser Ala Ala
        35                  40                  45

His Ser Pro Leu Asp Thr Ser Lys Gln Pro Leu Cys Gln Leu Trp Ala
    50                  55                  60

Glu Lys His Gly Ala Arg Gly Thr His Glu Val Arg Tyr Val Ser Ala
65                  70                  75                  80

Gly Gln Ser Val Ala Cys Gly Trp Trp Ala Phe Ala Pro Pro Cys Leu
                85                  90                  95

Gln Val Leu Asn Thr Pro Lys Gly Ile Leu Phe Phe Leu Cys Ala Ala
            100                 105                 110

Ala Phe Leu Gln Gly Met Thr Val Asn Gly Phe Ile Asn Thr Val Ile
        115                 120                 125

Thr Ser Leu Glu Arg Arg Tyr Asp Leu His Ser Tyr Gln Ser Gly Leu
    130                 135                 140

Ile Ala Ser Ser Tyr Asp Ile Ala Ala Cys Leu Cys Leu Thr Phe Val
145                 150                 155                 160

Ser Tyr Phe Gly Gly Ser Gly His Lys Pro Arg Trp Leu Gly Trp Gly
                165                 170                 175
```

-continued

```
Val Leu Leu Met Gly Thr Gly Ser Leu Val Phe Ala Leu Pro His Phe
            180                 185                 190

Thr Ala Gly Arg Tyr Glu Val Glu Leu Asp Ala Gly Val Arg Thr Cys
        195                 200                 205

Pro Ala Asn Pro Gly Ala Val Cys Ala Asp Ser Thr Ser Gly Leu Ser
    210                 215                 220

Arg Tyr Gln Leu Val Phe Met Leu Gly Gln Phe Leu His Gly Val Gly
225                 230                 235                 240

Ala Thr Pro Leu Tyr Thr Leu Gly Val Thr Tyr Leu Asp Glu Asn Val
                245                 250                 255

Lys Ser Ser Cys Ser Pro Val Tyr Ile Ala Ile Phe Tyr Thr Ala Ala
            260                 265                 270

Ile Leu Gly Pro Ala Ala Gly Tyr Leu Ile Gly Gly Ala Leu Leu Asn
        275                 280                 285

Ile Tyr Thr Glu Met Gly Arg Arg Thr Glu Leu Thr Thr Glu Ser Pro
    290                 295                 300

Leu Trp Val Gly Ala Trp Trp Val Gly Phe Leu Gly Ser Gly Ala Ala
305                 310                 315                 320

Ala Phe Phe Thr Ala Val Pro Ile Leu Gly Tyr Pro Arg Gln Leu Pro
                325                 330                 335

Gly Ser Gln Arg Tyr Ala Val Met Arg Ala Ala Glu Met His Gln Leu
            340                 345                 350

Lys Asp Ser Ser Arg Gly Glu Ala Ser Asn Pro Asp Phe Gly Lys Thr
        355                 360                 365

Ile Arg Asp Leu Pro Leu Ser Ile Trp Leu Leu Leu Lys Asn Pro Thr
    370                 375                 380

Phe Ile Leu Leu Cys Leu Ala Gly Ala Thr Glu Ala Thr Leu Ile Thr
385                 390                 395                 400

Gly Met Ser Thr Phe Ser Pro Lys Phe Leu Glu Ser Gln Phe Ser Leu
                405                 410                 415

Ser Ala Ser Glu Ala Ala Thr Leu Phe Gly Tyr Leu Val Val Pro Ala
            420                 425                 430

Gly Gly Gly Gly Thr Phe Leu Gly Gly Phe Phe Val Asn Lys Leu Arg
        435                 440                 445

Leu Arg Gly Ser Ala Val Ile Lys Phe Cys Leu Phe Cys Thr Val Val
    450                 455                 460

Ser Leu Leu Gly Ile Leu Val Phe Ser Leu His Cys Pro Ser Val Pro
465                 470                 475                 480

Met Ala Gly Val Thr Ala Ser Tyr Gly Gly Ser Leu Leu Pro Glu Gly
                485                 490                 495

His Leu Asn Leu Thr Ala Pro Cys Asn Ala Ala Cys Ser Cys Gln Pro
            500                 505                 510

Glu His Tyr Ser Pro Val Cys Gly Ser Asp Gly Leu Met Tyr Phe Ser
        515                 520                 525

Leu Cys His Ala Gly Cys Pro Ala Ala Thr Glu Thr Asn Val Asp Gly
    530                 535                 540

Gln Lys Val Tyr Arg Asp Cys Ser Cys Ile Pro Gln Asn Leu Ser Ser
545                 550                 555                 560

Gly Phe Gly His Ala Thr Ala Gly Lys Cys Thr Ser Thr Cys Gln Arg
                565                 570                 575

Lys Pro Leu Leu Leu Val Phe Ile Phe Val Val Ile Phe Phe Thr Phe
```

-continued

```
            580                 585                 590

Leu Ser Ser Ile Pro Ala Leu Thr Ala Thr Leu Arg Cys Val Arg Asp
        595                 600                 605

Pro Gln Arg Ser Phe Ala Leu Gly Ile Gln Trp Ile Val Val Arg Ile
    610                 615                 620

Leu Gly Gly Ile Pro Gly Pro Ile Ala Phe Gly Trp Val Ile Asp Lys
625                 630                 635                 640

Ala Cys Leu Leu Trp Gln Asp Gln Cys Gly Gln Gln Gly Ser Cys Leu
                645                 650                 655

Val Tyr Gln Asn Ser Ala Met Ser Arg Tyr Ile Leu Ile Met Gly Leu
            660                 665                 670

Leu Tyr Lys Val Leu Gly Val Leu Phe Phe Ala Ile Ala Cys Phe Leu
        675                 680                 685

Tyr Lys Pro Leu Ser Glu Ser Ser Asp Gly Leu Glu Thr Cys Leu Pro
    690                 695                 700

Ser Gln Ser Ser Ala Pro Asp Ser Ala Thr Asp Ser Gln Leu Gln Ser
705                 710                 715                 720

Ser Val


<210> SEQ ID NO 50
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (1)...(1)
<223> OTHER INFORMATION: Xaa at position 1 can be R or K
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (2)...(2)
<223> OTHER INFORMATION: Xaa at position 2 can be L or I
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (3)...(7)
<223> OTHER INFORMATION: Xaa = Any Amino Acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (8)...(8)
<223> OTHER INFORMATION: Xaa at position 8 can be H or Q

<400> SEQUENCE: 50

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu
 1               5


<210> SEQ ID NO 51
<211> LENGTH: 18
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (1)...(1)
<223> OTHER INFORMATION: Xaa at position 1 can be L, I, V, M, S, T, A
      or G
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (2)...(2)
<223> OTHER INFORMATION: Xaa at position 1 can be L, I, V, M, S, F, A
      or G
<220> FEATURE:
<221> NAME/KEY: VARIANT
```

-continued

<222> LOCATION: 3,4,7,12-17
<223> OTHER INFORMATION: Xaa at these positions can be any amino acid.
The length of the segment with any amino acid after
position 11 can be 4, 5, or 6 residues.
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (5)...(5)
<223> OTHER INFORMATION: Xaa at position 5 can be L, I, V, M, S, or A
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (6)...(6)
<223> OTHER INFORMATION: Xaa at position 6 can be D or E
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (8)...(8)
<223> OTHER INFORMATION: Xaa at position 8 can be L, I, V, M, F, Y, W
or A
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (11)...(11)
<223> OTHER INFORMATION: Xaa at position 11 can be R or K
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (18)...(18)
<223> OTHER INFORMATION: Xaa at position 18 can be G, S, T or A

<400> SEQUENCE: 51

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Arg Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa

<210> SEQ ID NO 52
<211> LENGTH: 26
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (1)...(1)
<223> OTHER INFORMATION: Xaa at position 1 can be L, I, V, M or F
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: 2,5,6,8-15,17,18,20-25
<223> OTHER INFORMATION: Xaa = Any Amino Acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (4)...(4)
<223> OTHER INFORMATION: Xaa at position 4 can be L,
I, V, M, F or A
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (16)...(16)
<223> OTHER INFORMATION: Xaa at position 16 can be L, I, F or Y
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (19)...(19)
<223> OTHER INFORMATION: Xaa at position 19 can be E or Q.
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (26)...(26)
<223> OTHER INFORMATION: Xaa at position 26 can be R or K

<400> SEQUENCE: 52

Xaa Xaa Gly Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25

<210> SEQ ID NO 53

-continued

```
<211> LENGTH: 3227
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (345)...(2483)

<400> SEQUENCE: 53

ccctttctcg ggaagcgcgc cattgtgttg gtacccggga attcgcggcc gcgtcgaccc      60

ccctgcggag ttgtgttttc tgggaatcag caacaaaytg caaagaaatg gctcaaaagc     120

ttcagctctt tctgtgccct gggagctgag atgcacgtca gtggccttgc cagcgtggcc     180

aattctctgc tgactgccag aaaaaagagg ccaggaagaa agaggaaaga gaagagatcg     240

ctcagggtct gtggtgtgtg gtccatcctc ttgctgagca cattgaaagg aactggctat     300

ctttgatctc ttcctccaga tcagagtcaa ggaatgtgtt tata atg gac act tca      356
Met Asp Thr Ser
 1

tcc aaa gaa aat atc cag ttg ttc tgc aaa act tca gtg caa cct gtt       404
Ser Lys Glu Asn Ile Gln Leu Phe Cys Lys Thr Ser Val Gln Pro Val
 5                  10                  15                  20

gga agg cct tct ttt aaa aca gaa tat ccc tcc tca gaa gaa aag caa       452
Gly Arg Pro Ser Phe Lys Thr Glu Tyr Pro Ser Ser Glu Glu Lys Gln
                25                  30                  35

cca tgc tgt ggt gaa cta aag gtg ttc ttg tgt gcc ttg tct ttt gtt       500
Pro Cys Cys Gly Glu Leu Lys Val Phe Leu Cys Ala Leu Ser Phe Val
            40                  45                  50

tac ttt gcc aaa gca ttg gca gaa ggc tat ctg aag agc acc atc act       548
Tyr Phe Ala Lys Ala Leu Ala Glu Gly Tyr Leu Lys Ser Thr Ile Thr
        55                  60                  65

cag ata gag aga agg ttt gat atc cct tct tca ctg gtg gga gtt att       596
Gln Ile Glu Arg Arg Phe Asp Ile Pro Ser Ser Leu Val Gly Val Ile
    70                  75                  80

gat ggt agt ttt gaa att ggg aat ctc tta gtt ata aca ttt gtt agc       644
Asp Gly Ser Phe Glu Ile Gly Asn Leu Leu Val Ile Thr Phe Val Ser
 85                  90                  95                 100

tac ttt gga gcc aaa ctt cac agg cca aaa ata att gga gca ggg tgt       692
Tyr Phe Gly Ala Lys Leu His Arg Pro Lys Ile Ile Gly Ala Gly Cys
                105                 110                 115

gta atc atg gga gtt gga aca ctg ctc att gca atg cct cag ttc ttc       740
Val Ile Met Gly Val Gly Thr Leu Leu Ile Ala Met Pro Gln Phe Phe
            120                 125                 130

atg gag cag tac aaa tat gag aga tat tct cct tcc tcc aat tct act       788
Met Glu Gln Tyr Lys Tyr Glu Arg Tyr Ser Pro Ser Ser Asn Ser Thr
        135                 140                 145

ctc agc atc tct ccg tgt ctc cta gag tca agc agt caa tta cca gtt       836
Leu Ser Ile Ser Pro Cys Leu Leu Glu Ser Ser Ser Gln Leu Pro Val
    150                 155                 160

tca gtt atg gaa aaa tca aaa tcc aaa ata agt aac gaa tgt gaa gtg       884
Ser Val Met Glu Lys Ser Lys Ser Lys Ile Ser Asn Glu Cys Glu Val
165                 170                 175                 180

gac act agc tct tcc atg tgg att tat gtt ttc ctg ggc aat ctt ctt       932
Asp Thr Ser Ser Ser Met Trp Ile Tyr Val Phe Leu Gly Asn Leu Leu
                185                 190                 195

cgt gga ata gga gaa act ccc att cag cct ttg ggc att gcc tac ctg       980
Arg Gly Ile Gly Glu Thr Pro Ile Gln Pro Leu Gly Ile Ala Tyr Leu
            200                 205                 210

gat gat ttt gcc agt gaa gac aat gca gct ttc tat att ggg tgt gtg      1028
```

-continued

```
Asp Asp Phe Ala Ser Glu Asp Asn Ala Ala Phe Tyr Ile Gly Cys Val
        215                 220                 225

cag acg gtt gca att ata gga cca atc ttt ggt ttc ctg tta ggc tca     1076
Gln Thr Val Ala Ile Ile Gly Pro Ile Phe Gly Phe Leu Leu Gly Ser
    230                 235                 240

tta tgt gcc aaa cta tat gtt gac att ggc ttt gta aac cta gat cac     1124
Leu Cys Ala Lys Leu Tyr Val Asp Ile Gly Phe Val Asn Leu Asp His
245                 250                 255                 260

ata acc att acc cca aaa gat ccc cag tgg gta gga gcc tgg tgg ctt     1172
Ile Thr Ile Thr Pro Lys Asp Pro Gln Trp Val Gly Ala Trp Trp Leu
                265                 270                 275

ggc tat cta ata gca gga atc ata agt ctt ctt gca gct gtg cct ttc     1220
Gly Tyr Leu Ile Ala Gly Ile Ile Ser Leu Leu Ala Ala Val Pro Phe
            280                 285                 290

tgg tat tta cca aag agt tta cca aga tcc caa agt aga gag gat tct     1268
Trp Tyr Leu Pro Lys Ser Leu Pro Arg Ser Gln Ser Arg Glu Asp Ser
        295                 300                 305

aat tct tcc tct gag aaa tcc aag ttt att ata gat gat cac aca gac     1316
Asn Ser Ser Ser Glu Lys Ser Lys Phe Ile Ile Asp Asp His Thr Asp
    310                 315                 320

tac caa aca ccc cag gga gaa aat gca aaa ata atg gaa atg gca aga     1364
Tyr Gln Thr Pro Gln Gly Glu Asn Ala Lys Ile Met Glu Met Ala Arg
325                 330                 335                 340

gat ttt ctt cca tca ctg aag aat ctt ttt gga aac cca gta tac ttc     1412
Asp Phe Leu Pro Ser Leu Lys Asn Leu Phe Gly Asn Pro Val Tyr Phe
                345                 350                 355

cta tat tta tgt aca agc act gtt cag ttc aat tct ctg ttc ggc atg     1460
Leu Tyr Leu Cys Thr Ser Thr Val Gln Phe Asn Ser Leu Phe Gly Met
            360                 365                 370

gtg acg tac aaa cca aag tac att gag cag cag tat gga cag tca tcc     1508
Val Thr Tyr Lys Pro Lys Tyr Ile Glu Gln Gln Tyr Gly Gln Ser Ser
        375                 380                 385

tcc agg gcc aac ttt gtg atc ggg ctc atc aac att cca gca gtg gcc     1556
Ser Arg Ala Asn Phe Val Ile Gly Leu Ile Asn Ile Pro Ala Val Ala
    390                 395                 400

ctt gga ata ttc tct ggg ggg ata gtt atg aaa aaa ttc aga atc agt     1604
Leu Gly Ile Phe Ser Gly Gly Ile Val Met Lys Lys Phe Arg Ile Ser
405                 410                 415                 420

gtg tgt gga gct gca aaa ctc tac ttg gga tca tct gtc ttt ggt tac     1652
Val Cys Gly Ala Ala Lys Leu Tyr Leu Gly Ser Ser Val Phe Gly Tyr
                425                 430                 435

ctc cta ttt ctt tcc ctg ttt gca ctg ggc tgt gaa aat tct gat gtg     1700
Leu Leu Phe Leu Ser Leu Phe Ala Leu Gly Cys Glu Asn Ser Asp Val
            440                 445                 450

gca gga cta act gtc tcc tac caa gga acc aaa cct gtc tct tat cat     1748
Ala Gly Leu Thr Val Ser Tyr Gln Gly Thr Lys Pro Val Ser Tyr His
        455                 460                 465

gaa cga gct ctc ttt tca gat tgc aac tca aga tgc aaa tgt tca gag     1796
Glu Arg Ala Leu Phe Ser Asp Cys Asn Ser Arg Cys Lys Cys Ser Glu
    470                 475                 480

aca aaa tgg gaa ccc atg tgc ggt gaa aat gga atc aca tat gta tca     1844
Thr Lys Trp Glu Pro Met Cys Gly Glu Asn Gly Ile Thr Tyr Val Ser
485                 490                 495                 500

gct tgt ctt gct ggt tgt caa acc tcc aac agg agt gga aaa aat att     1892
Ala Cys Leu Ala Gly Cys Gln Thr Ser Asn Arg Ser Gly Lys Asn Ile
                505                 510                 515

ata ttt tac aac tgc act tgt gtg gga att gca gct tct aaa tcc gga     1940
```

-continued

```
Ile Phe Tyr Asn Cys Thr Cys Val Gly Ile Ala Ala Ser Lys Ser Gly
            520                 525                 530

aat tcc tca ggc ata gtg gga aga tgt cag aaa gac aat gga tgt ccc     1988
Asn Ser Ser Gly Ile Val Gly Arg Cys Gln Lys Asp Asn Gly Cys Pro
        535                 540                 545

caa atg ttt ctg tat ttc ctt gta att tca gtc atc aca tcc tat act     2036
Gln Met Phe Leu Tyr Phe Leu Val Ile Ser Val Ile Thr Ser Tyr Thr
    550                 555                 560

tta tcc cta ggt ggc ata cct gga tac ata tta ctt ctg agg tgc att     2084
Leu Ser Leu Gly Gly Ile Pro Gly Tyr Ile Leu Leu Leu Arg Cys Ile
565                 570                 575                 580

aag cca cag ctt aag tct ttt gcc ttg ggt atc tac aca tta gca ata     2132
Lys Pro Gln Leu Lys Ser Phe Ala Leu Gly Ile Tyr Thr Leu Ala Ile
                585                 590                 595

aga gtt ctt gca gga atc cca gct cca gtg tat ttt gga gtt ttg att     2180
Arg Val Leu Ala Gly Ile Pro Ala Pro Val Tyr Phe Gly Val Leu Ile
            600                 605                 610

gat act tca tgc ctc aaa tgg gga ttt aaa aga tgt gga agt aga gga     2228
Asp Thr Ser Cys Leu Lys Trp Gly Phe Lys Arg Cys Gly Ser Arg Gly
        615                 620                 625

tca tgc aga tta tat gat tca aat gtc ttc aga cat ata tat ctg gga     2276
Ser Cys Arg Leu Tyr Asp Ser Asn Val Phe Arg His Ile Tyr Leu Gly
    630                 635                 640

cta act gtg ata ctg ggc aca gtg tca att ctc cta agc att gca gta     2324
Leu Thr Val Ile Leu Gly Thr Val Ser Ile Leu Leu Ser Ile Ala Val
645                 650                 655                 660

ctt ttc att tta aag aaa aat tat gtt tca aaa cac aga agt ttt ata     2372
Leu Phe Ile Leu Lys Lys Asn Tyr Val Ser Lys His Arg Ser Phe Ile
                665                 670                 675

acc aag aga gaa aga aca atg gtg tct aca aga ttc caa aag gaa aat     2420
Thr Lys Arg Glu Arg Thr Met Val Ser Thr Arg Phe Gln Lys Glu Asn
            680                 685                 690

tac act aca agt gat cat ctg cta caa ccc aac tac tgg cca ggc aag     2468
Tyr Thr Thr Ser Asp His Leu Leu Gln Pro Asn Tyr Trp Pro Gly Lys
        695                 700                 705

gaa act caa ctt tag aaacatgatg actggaagtc atgtcttcta attggttgac     2523
Glu Thr Gln Leu  *
    710

attttgcaaa caaataaatt gtaatcaaaa gagctctaaa tttgtaattt ctttctcctt   2583

tcaaaaaatg tctactttgt tttggtccta ggcattaggt aatataactg ataatatact   2643

gaaacatata atggaagatg cagatgataa aactaatttt gaacttttta atttatataa   2703

attattttat atcacttact tatttcactt tattttgctt tgtgctcatt gatatatatt   2763

agctgtactc ctagaagaac aattgtctct attgtcacac atggttatat ttaaagtaat   2823

ttctgaactg tgtaatgtgt ctagagtaag caaatactgc taacaattaa ctcatacctt   2883

gggttccttc aagtattact cctatagtat tttctcccat agctgtcttc atctgtgtat   2943

tttaataatg atcttaggat ggagcagaac atggagagga agatttcatt ttaagctcct   3003

ccttttcttt gaaatacaat aatttatata gaaatgtgta gcagcaaatt atattgggga   3063

ttagaatttt gaattaatag ctctcctact attaatttac atgtgctttt tgtgtggcgc   3123

tataagtgac tatggttgta aagtaataaa attgatgtta acatgcccaa aaaaaaaaaa   3183

aaaaaaaacc aaaaaaaaaa aaaaaaaagg gcgggccgct agac                    3227
```

-continued

<210> SEQ ID NO 54
<211> LENGTH: 712
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 54

Met Asp Thr Ser Ser Lys Glu Asn Ile Gln Leu Phe Cys Lys Thr Ser
1 5 10 15

Val Gln Pro Val Gly Arg Pro Ser Phe Lys Thr Glu Tyr Pro Ser Ser
20 25 30

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

Leu Ser Phe Val Tyr Phe Ala Lys Ala Leu Ala Glu Gly Tyr Leu Lys
50 55 60

Ser Thr Ile Thr Gln Ile Glu Arg Arg Phe Asp Ile Pro Ser Ser Leu
65 70 75 80

Val Gly Val Ile Asp Gly Ser Phe Glu Ile Gly Asn Leu Leu Val Ile
85 90 95

Thr Phe Val Ser Tyr Phe Gly Ala Lys Leu His Arg Pro Lys Ile Ile
100 105 110

Gly Ala Gly Cys Val Ile Met Gly Val Gly Thr Leu Leu Ile Ala Met
115 120 125

Pro Gln Phe Phe Met Glu Gln Tyr Lys Tyr Glu Arg Tyr Ser Pro Ser
130 135 140

Ser Asn Ser Thr Leu Ser Ile Ser Pro Cys Leu Leu Glu Ser Ser Ser
145 150 155 160

Gln Leu Pro Val Ser Val Met Glu Lys Ser Lys Ser Lys Ile Ser Asn
165 170 175

Glu Cys Glu Val Asp Thr Ser Ser Ser Met Trp Ile Tyr Val Phe Leu
180 185 190

Gly Asn Leu Leu Arg Gly Ile Gly Glu Thr Pro Ile Gln Pro Leu Gly
195 200 205

Ile Ala Tyr Leu Asp Asp Phe Ala Ser Glu Asp Asn Ala Ala Phe Tyr
210 215 220

Ile Gly Cys Val Gln Thr Val Ala Ile Ile Gly Pro Ile Phe Gly Phe
225 230 235 240

Leu Leu Gly Ser Leu Cys Ala Lys Leu Tyr Val Asp Ile Gly Phe Val
245 250 255

Asn Leu Asp His Ile Thr Ile Thr Pro Lys Asp Pro Gln Trp Val Gly
260 265 270

Ala Trp Trp Leu Gly Tyr Leu Ile Ala Gly Ile Ile Ser Leu Leu Ala
275 280 285

Ala Val Pro Phe Trp Tyr Leu Pro Lys Ser Leu Pro Arg Ser Gln Ser
290 295 300

Arg Glu Asp Ser Asn Ser Ser Ser Glu Lys Ser Lys Phe Ile Ile Asp
305 310 315 320

Asp His Thr Asp Tyr Gln Thr Pro Gln Gly Glu Asn Ala Lys Ile Met
325 330 335

Glu Met Ala Arg Asp Phe Leu Pro Ser Leu Lys Asn Leu Phe Gly Asn
340 345 350

Pro Val Tyr Phe Leu Tyr Leu Cys Thr Ser Thr Val Gln Phe Asn Ser
355 360 365

-continued

```
Leu Phe Gly Met Val Thr Tyr Lys Pro Lys Tyr Ile Glu Gln Gln Tyr
    370                 375                 380

Gly Gln Ser Ser Ser Arg Ala Asn Phe Val Ile Gly Leu Ile Asn Ile
385                 390                 395                 400

Pro Ala Val Ala Leu Gly Ile Phe Ser Gly Gly Ile Val Met Lys Lys
                405                 410                 415

Phe Arg Ile Ser Val Cys Gly Ala Ala Lys Leu Tyr Leu Gly Ser Ser
            420                 425                 430

Val Phe Gly Tyr Leu Leu Phe Leu Ser Leu Phe Ala Leu Gly Cys Glu
        435                 440                 445

Asn Ser Asp Val Ala Gly Leu Thr Val Ser Tyr Gln Gly Thr Lys Pro
    450                 455                 460

Val Ser Tyr His Glu Arg Ala Leu Phe Ser Asp Cys Asn Ser Arg Cys
465                 470                 475                 480

Lys Cys Ser Glu Thr Lys Trp Glu Pro Met Cys Gly Glu Asn Gly Ile
                485                 490                 495

Thr Tyr Val Ser Ala Cys Leu Ala Gly Cys Gln Thr Ser Asn Arg Ser
            500                 505                 510

Gly Lys Asn Ile Ile Phe Tyr Asn Cys Thr Cys Val Gly Ile Ala Ala
        515                 520                 525

Ser Lys Ser Gly Asn Ser Ser Gly Ile Val Gly Arg Cys Gln Lys Asp
    530                 535                 540

Asn Gly Cys Pro Gln Met Phe Leu Tyr Phe Leu Val Ile Ser Val Ile
545                 550                 555                 560

Thr Ser Tyr Thr Leu Ser Leu Gly Gly Ile Pro Gly Tyr Ile Leu Leu
                565                 570                 575

Leu Arg Cys Ile Lys Pro Gln Leu Lys Ser Phe Ala Leu Gly Ile Tyr
            580                 585                 590

Thr Leu Ala Ile Arg Val Leu Ala Gly Ile Pro Ala Pro Val Tyr Phe
        595                 600                 605

Gly Val Leu Ile Asp Thr Ser Cys Leu Lys Trp Gly Phe Lys Arg Cys
    610                 615                 620

Gly Ser Arg Gly Ser Cys Arg Leu Tyr Asp Ser Asn Val Phe Arg His
625                 630                 635                 640

Ile Tyr Leu Gly Leu Thr Val Ile Leu Gly Thr Val Ser Ile Leu Leu
                645                 650                 655

Ser Ile Ala Val Leu Phe Ile Leu Lys Lys Asn Tyr Val Ser Lys His
            660                 665                 670

Arg Ser Phe Ile Thr Lys Arg Glu Arg Thr Met Val Ser Thr Arg Phe
        675                 680                 685

Gln Lys Glu Asn Tyr Thr Thr Ser Asp His Leu Leu Gln Pro Asn Tyr
    690                 695                 700

Trp Pro Gly Lys Glu Thr Gln Leu
705                 710


<210> SEQ ID NO 55
<211> LENGTH: 2139
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(2139)

<400> SEQUENCE: 55
```

-continued

```
atg gac act tca tcc aaa gaa aat atc cag ttg ttc tgc aaa act tca       48
Met Asp Thr Ser Ser Lys Glu Asn Ile Gln Leu Phe Cys Lys Thr Ser
 1               5                  10                  15

gtg caa cct gtt gga agg cct tct ttt aaa aca gaa tat ccc tcc tca       96
Val Gln Pro Val Gly Arg Pro Ser Phe Lys Thr Glu Tyr Pro Ser Ser
            20                  25                  30

gaa gaa aag caa cca tgc tgt ggt gaa cta aag gtg ttc ttg tgt gcc      144
Glu Glu Lys Gln Pro Cys Cys Gly Glu Leu Lys Val Phe Leu Cys Ala
        35                  40                  45

ttg tct ttt gtt tac ttt gcc aaa gca ttg gca gaa ggc tat ctg aag      192
Leu Ser Phe Val Tyr Phe Ala Lys Ala Leu Ala Glu Gly Tyr Leu Lys
    50                  55                  60

agc acc atc act cag ata gag aga agg ttt gat atc cct tct tca ctg      240
Ser Thr Ile Thr Gln Ile Glu Arg Arg Phe Asp Ile Pro Ser Ser Leu
 65                  70                  75                  80

gtg gga gtt att gat ggt agt ttt gaa att ggg aat ctc tta gtt ata      288
Val Gly Val Ile Asp Gly Ser Phe Glu Ile Gly Asn Leu Leu Val Ile
                85                  90                  95

aca ttt gtt agc tac ttt gga gcc aaa ctt cac agg cca aaa ata att      336
Thr Phe Val Ser Tyr Phe Gly Ala Lys Leu His Arg Pro Lys Ile Ile
            100                 105                 110

gga gca ggg tgt gta atc atg gga gtt gga aca ctg ctc att gca atg      384
Gly Ala Gly Cys Val Ile Met Gly Val Gly Thr Leu Leu Ile Ala Met
        115                 120                 125

cct cag ttc ttc atg gag cag tac aaa tat gag aga tat tct cct tcc      432
Pro Gln Phe Phe Met Glu Gln Tyr Lys Tyr Glu Arg Tyr Ser Pro Ser
    130                 135                 140

tcc aat tct act ctc agc atc tct ccg tgt ctc cta gag tca agc agt      480
Ser Asn Ser Thr Leu Ser Ile Ser Pro Cys Leu Leu Glu Ser Ser Ser
145                 150                 155                 160

caa tta cca gtt tca gtt atg gaa aaa tca aaa tcc aaa ata agt aac      528
Gln Leu Pro Val Ser Val Met Glu Lys Ser Lys Ser Lys Ile Ser Asn
                165                 170                 175

gaa tgt gaa gtg gac act agc tct tcc atg tgg att tat gtt ttc ctg      576
Glu Cys Glu Val Asp Thr Ser Ser Ser Met Trp Ile Tyr Val Phe Leu
            180                 185                 190

ggc aat ctt ctt cgt gga ata gga gaa act ccc att cag cct ttg ggc      624
Gly Asn Leu Leu Arg Gly Ile Gly Glu Thr Pro Ile Gln Pro Leu Gly
        195                 200                 205

att gcc tac ctg gat gat ttt gcc agt gaa gac aat gca gct ttc tat      672
Ile Ala Tyr Leu Asp Asp Phe Ala Ser Glu Asp Asn Ala Ala Phe Tyr
    210                 215                 220

att ggg tgt gtg cag acg gtt gca att ata gga cca atc ttt ggt ttc      720
Ile Gly Cys Val Gln Thr Val Ala Ile Ile Gly Pro Ile Phe Gly Phe
225                 230                 235                 240

ctg tta ggc tca tta tgt gcc aaa cta tat gtt gac att ggc ttt gta      768
Leu Leu Gly Ser Leu Cys Ala Lys Leu Tyr Val Asp Ile Gly Phe Val
                245                 250                 255

aac cta gat cac ata acc att acc cca aaa gat ccc cag tgg gta gga      816
Asn Leu Asp His Ile Thr Ile Thr Pro Lys Asp Pro Gln Trp Val Gly
            260                 265                 270

gcc tgg tgg ctt ggc tat cta ata gca gga atc ata agt ctt ctt gca      864
Ala Trp Trp Leu Gly Tyr Leu Ile Ala Gly Ile Ile Ser Leu Leu Ala
        275                 280                 285

gct gtg cct ttc tgg tat tta cca aag agt tta cca aga tcc caa agt      912
Ala Val Pro Phe Trp Tyr Leu Pro Lys Ser Leu Pro Arg Ser Gln Ser
    290                 295                 300
```

-continued

```
aga gag gat tct aat tct tcc tct gag aaa tcc aag ttt att ata gat      960
Arg Glu Asp Ser Asn Ser Ser Ser Glu Lys Ser Lys Phe Ile Ile Asp
305                 310                 315                 320

gat cac aca gac tac caa aca ccc cag gga gaa aat gca aaa ata atg     1008
Asp His Thr Asp Tyr Gln Thr Pro Gln Gly Glu Asn Ala Lys Ile Met
                325                 330                 335

gaa atg gca aga gat ttt ctt cca tca ctg aag aat ctt ttt gga aac     1056
Glu Met Ala Arg Asp Phe Leu Pro Ser Leu Lys Asn Leu Phe Gly Asn
            340                 345                 350

cca gta tac ttc cta tat tta tgt aca agc act gtt cag ttc aat tct     1104
Pro Val Tyr Phe Leu Tyr Leu Cys Thr Ser Thr Val Gln Phe Asn Ser
        355                 360                 365

ctg ttc ggc atg gtg acg tac aaa cca aag tac att gag cag cag tat     1152
Leu Phe Gly Met Val Thr Tyr Lys Pro Lys Tyr Ile Glu Gln Gln Tyr
    370                 375                 380

gga cag tca tcc tcc agg gcc aac ttt gtg atc ggg ctc atc aac att     1200
Gly Gln Ser Ser Ser Arg Ala Asn Phe Val Ile Gly Leu Ile Asn Ile
385                 390                 395                 400

cca gca gtg gcc ctt gga ata ttc tct ggg ggg ata gtt atg aaa aaa     1248
Pro Ala Val Ala Leu Gly Ile Phe Ser Gly Gly Ile Val Met Lys Lys
                405                 410                 415

ttc aga atc agt gtg tgt gga gct gca aaa ctc tac ttg gga tca tct     1296
Phe Arg Ile Ser Val Cys Gly Ala Ala Lys Leu Tyr Leu Gly Ser Ser
            420                 425                 430

gtc ttt ggt tac ctc cta ttt ctt tcc ctg ttt gca ctg ggc tgt gaa     1344
Val Phe Gly Tyr Leu Leu Phe Leu Ser Leu Phe Ala Leu Gly Cys Glu
        435                 440                 445

aat tct gat gtg gca gga cta act gtc tcc tac caa gga acc aaa cct     1392
Asn Ser Asp Val Ala Gly Leu Thr Val Ser Tyr Gln Gly Thr Lys Pro
    450                 455                 460

gtc tct tat cat gaa cga gct ctc ttt tca gat tgc aac tca aga tgc     1440
Val Ser Tyr His Glu Arg Ala Leu Phe Ser Asp Cys Asn Ser Arg Cys
465                 470                 475                 480

aaa tgt tca gag aca aaa tgg gaa ccc atg tgc ggt gaa aat gga atc     1488
Lys Cys Ser Glu Thr Lys Trp Glu Pro Met Cys Gly Glu Asn Gly Ile
                485                 490                 495

aca tat gta tca gct tgt ctt gct ggt tgt caa acc tcc aac agg agt     1536
Thr Tyr Val Ser Ala Cys Leu Ala Gly Cys Gln Thr Ser Asn Arg Ser
            500                 505                 510

gga aaa aat att ata ttt tac aac tgc act tgt gtg gga att gca gct     1584
Gly Lys Asn Ile Ile Phe Tyr Asn Cys Thr Cys Val Gly Ile Ala Ala
        515                 520                 525

tct aaa tcc gga aat tcc tca ggc ata gtg gga aga tgt cag aaa gac     1632
Ser Lys Ser Gly Asn Ser Ser Gly Ile Val Gly Arg Cys Gln Lys Asp
    530                 535                 540

aat gga tgt ccc caa atg ttt ctg tat ttc ctt gta att tca gtc atc     1680
Asn Gly Cys Pro Gln Met Phe Leu Tyr Phe Leu Val Ile Ser Val Ile
545                 550                 555                 560

aca tcc tat act tta tcc cta ggt ggc ata cct gga tac ata tta ctt     1728
Thr Ser Tyr Thr Leu Ser Leu Gly Gly Ile Pro Gly Tyr Ile Leu Leu
                565                 570                 575

ctg agg tgc att aag cca cag ctt aag tct ttt gcc ttg ggt atc tac     1776
Leu Arg Cys Ile Lys Pro Gln Leu Lys Ser Phe Ala Leu Gly Ile Tyr
            580                 585                 590

aca tta gca ata aga gtt ctt gca gga atc cca gct cca gtg tat ttt     1824
Thr Leu Ala Ile Arg Val Leu Ala Gly Ile Pro Ala Pro Val Tyr Phe
        595                 600                 605
```

-continued

```
gga gtt ttg att gat act tca tgc ctc aaa tgg gga ttt aaa aga tgt     1872
Gly Val Leu Ile Asp Thr Ser Cys Leu Lys Trp Gly Phe Lys Arg Cys
    610                 615                 620

gga agt aga gga tca tgc aga tta tat gat tca aat gtc ttc aga cat     1920
Gly Ser Arg Gly Ser Cys Arg Leu Tyr Asp Ser Asn Val Phe Arg His
625                 630                 635                 640

ata tat ctg gga cta act gtg ata ctg ggc aca gtg tca att ctc cta     1968
Ile Tyr Leu Gly Leu Thr Val Ile Leu Gly Thr Val Ser Ile Leu Leu
                645                 650                 655

agc att gca gta ctt ttc att tta aag aaa aat tat gtt tca aaa cac     2016
Ser Ile Ala Val Leu Phe Ile Leu Lys Lys Asn Tyr Val Ser Lys His
            660                 665                 670

aga agt ttt ata acc aag aga gaa aga aca atg gtg tct aca aga ttc     2064
Arg Ser Phe Ile Thr Lys Arg Glu Arg Thr Met Val Ser Thr Arg Phe
        675                 680                 685

caa aag gaa aat tac act aca agt gat cat ctg cta caa ccc aac tac     2112
Gln Lys Glu Asn Tyr Thr Thr Ser Asp His Leu Leu Gln Pro Asn Tyr
    690                 695                 700

tgg cca ggc aag gaa act caa ctt tag                                 2139
Trp Pro Gly Lys Glu Thr Gln Leu  *
705                 710


<210> SEQ ID NO 56
<211> LENGTH: 4197
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(3348)

<400> SEQUENCE: 56

atg agg aga ctg agt ttg tgg tgg ctg ctg agc agg gtc tgt ctg ctg       48
Met Arg Arg Leu Ser Leu Trp Trp Leu Leu Ser Arg Val Cys Leu Leu
 1               5                  10                  15

ttg ccg ccg ccc tgc gca ctg gtg ctg gcc ggg gtg ccc agc tcc tcc       96
Leu Pro Pro Pro Cys Ala Leu Val Leu Ala Gly Val Pro Ser Ser Ser
            20                  25                  30

tcg cac ccg cag ccc tgc cag atc ctc aag cgc atc ggg cac gcg gtg      144
Ser His Pro Gln Pro Cys Gln Ile Leu Lys Arg Ile Gly His Ala Val
        35                  40                  45

agg gtg ggc gcg gtg cac ttg cag ccc tgg acc acc gcc ccc cgc gcg      192
Arg Val Gly Ala Val His Leu Gln Pro Trp Thr Thr Ala Pro Arg Ala
    50                  55                  60

gcc agc cgc gct ccg gac gac agc cga gca gga gcc cag agg gat gag      240
Ala Ser Arg Ala Pro Asp Asp Ser Arg Ala Gly Ala Gln Arg Asp Glu
65                  70                  75                  80

ccg gag cca ggg act agg cgg tcc ccg gcg ccc tcg ccg ggc gca cgc      288
Pro Glu Pro Gly Thr Arg Arg Ser Pro Ala Pro Ser Pro Gly Ala Arg
                85                  90                  95

tgg ttg ggg agc acc ctg cat ggc cgg ggg ccg ccg ggc tcc cgt aag      336
Trp Leu Gly Ser Thr Leu His Gly Arg Gly Pro Pro Gly Ser Arg Lys
            100                 105                 110

ccc ggg gag ggc gcc agg gcg gag gcc ctg tgg cca cgg gac gcc ctc      384
Pro Gly Glu Gly Ala Arg Ala Glu Ala Leu Trp Pro Arg Asp Ala Leu
        115                 120                 125

cta ttt gcc gtg gac aac ctg aac cgc gtg gaa ggg ctg cta ccc tac      432
Leu Phe Ala Val Asp Asn Leu Asn Arg Val Glu Gly Leu Leu Pro Tyr
    130                 135                 140
```

-continued

```
aac ctg tct ttg gaa gta gtg atg gcc atc gag gca ggc ctg ggc gat      480
Asn Leu Ser Leu Glu Val Val Met Ala Ile Glu Ala Gly Leu Gly Asp
145                 150                 155                 160

ctg cca ctt ttg ccc ttc tcc tcc cct agt tcg cca tgg agc agt gac      528
Leu Pro Leu Leu Pro Phe Ser Ser Pro Ser Ser Pro Trp Ser Ser Asp
                165                 170                 175

cct ttc tcc ttc ctg caa agt gtg tgc cat acc gtg gtg gtg caa ggg      576
Pro Phe Ser Phe Leu Gln Ser Val Cys His Thr Val Val Val Gln Gly
            180                 185                 190

gtg tcg gcg ctg ctc gcc ttc ccc cag agc cag ggc gaa atg atg gag      624
Val Ser Ala Leu Leu Ala Phe Pro Gln Ser Gln Gly Glu Met Met Glu
        195                 200                 205

ctc gac ttg gtc agc tta gtc ctg cac att cca gtg atc agc atc gtg      672
Leu Asp Leu Val Ser Leu Val Leu His Ile Pro Val Ile Ser Ile Val
    210                 215                 220

cgc cac gag ttt ccg cgg gag agt cag aat ccc ctt cac cta caa ctg      720
Arg His Glu Phe Pro Arg Glu Ser Gln Asn Pro Leu His Leu Gln Leu
225                 230                 235                 240

agt tta gaa aat tca tta agt tct gat gct gat gtc act gtc tca atc      768
Ser Leu Glu Asn Ser Leu Ser Ser Asp Ala Asp Val Thr Val Ser Ile
                245                 250                 255

ctg acc atg aac aac tgg tac aat ttt agc ttg ttg ctg tgc cag gaa      816
Leu Thr Met Asn Asn Trp Tyr Asn Phe Ser Leu Leu Leu Cys Gln Glu
            260                 265                 270

gac tgg aac atc acc gac ttc ctc ctc ctt acc cag aat aat tcc aag      864
Asp Trp Asn Ile Thr Asp Phe Leu Leu Leu Thr Gln Asn Asn Ser Lys
        275                 280                 285

ttc cac ctt ggt tct atc atc aac atc acc gct aac ctc ccc tcc acc      912
Phe His Leu Gly Ser Ile Ile Asn Ile Thr Ala Asn Leu Pro Ser Thr
    290                 295                 300

cag gac ctc ttg agc ttc cta cag atc cag ctt gag agt att aag aac      960
Gln Asp Leu Leu Ser Phe Leu Gln Ile Gln Leu Glu Ser Ile Lys Asn
305                 310                 315                 320

agc aca ccc aca gtg gtg atg ttt ggc tgc gac atg gaa agt atc cgg     1008
Ser Thr Pro Thr Val Val Met Phe Gly Cys Asp Met Glu Ser Ile Arg
                325                 330                 335

cgg att ttc gaa att aca acc cag ttt ggg gtc atg ccc cct gaa ctt     1056
Arg Ile Phe Glu Ile Thr Thr Gln Phe Gly Val Met Pro Pro Glu Leu
            340                 345                 350

cgt tgg gtg ctg gga gat tcc cag aat atg gag gaa ctg agg aca gag     1104
Arg Trp Val Leu Gly Asp Ser Gln Asn Met Glu Glu Leu Arg Thr Glu
        355                 360                 365

ggt ctg ccc tta gga ctc att gct cat gga aaa aca aca cag tct gtc     1152
Gly Leu Pro Leu Gly Leu Ile Ala His Gly Lys Thr Thr Gln Ser Val
    370                 375                 380

ttt gag cac tac gta caa gat gct atg gag ctg gtc gca aga gct gta     1200
Phe Glu His Tyr Val Gln Asp Ala Met Glu Leu Val Ala Arg Ala Val
385                 390                 395                 400

gcc aca gcc acc atg atc caa cca gaa ctt gct ctc att ccc agc acg     1248
Ala Thr Ala Thr Met Ile Gln Pro Glu Leu Ala Leu Ile Pro Ser Thr
                405                 410                 415

atg aac tgc atg gag gtg gaa act aca aat ctc act tca gga caa tat     1296
Met Asn Cys Met Glu Val Glu Thr Thr Asn Leu Thr Ser Gly Gln Tyr
            420                 425                 430

tta tca agg ttt cta gcc aat acc act ttc aga ggc ctc agt ggt tcc     1344
Leu Ser Arg Phe Leu Ala Asn Thr Thr Phe Arg Gly Leu Ser Gly Ser
        435                 440                 445
```

-continued

```
att aga gta aaa ggt tcc acc atc gtc agc tca gaa aac aac ttt ttc     1392
Ile Arg Val Lys Gly Ser Thr Ile Val Ser Ser Glu Asn Asn Phe Phe
    450                 455                 460

atc tgg aat ctt caa cat gac ccc atg gga aag cca atg tgg acc cgc     1440
Ile Trp Asn Leu Gln His Asp Pro Met Gly Lys Pro Met Trp Thr Arg
465                 470                 475                 480

ttg ggc agc tgg cag ggg aga aag att gtc atg gac tat gga ata tgg     1488
Leu Gly Ser Trp Gln Gly Arg Lys Ile Val Met Asp Tyr Gly Ile Trp
                485                 490                 495

cca gag cag gcc cag aga cac aaa acc cac ttc caa cat cca agt aag     1536
Pro Glu Gln Ala Gln Arg His Lys Thr His Phe Gln His Pro Ser Lys
            500                 505                 510

cta cac ttg aga gtg gtt acc ctg att gag cat cct ttt gtc ttc aca     1584
Leu His Leu Arg Val Val Thr Leu Ile Glu His Pro Phe Val Phe Thr
        515                 520                 525

agg gag gta gat gat gaa ggc ttg tgc cct gct ggc caa ctc tgt cta     1632
Arg Glu Val Asp Asp Glu Gly Leu Cys Pro Ala Gly Gln Leu Cys Leu
    530                 535                 540

gac ccc atg act aat gac tct tcc aca ctg gac agc ctt ttt agc agc     1680
Asp Pro Met Thr Asn Asp Ser Ser Thr Leu Asp Ser Leu Phe Ser Ser
545                 550                 555                 560

ctc cat agc agt aat gat aca gtg ccc att aaa ttc aag aag tgc tgc     1728
Leu His Ser Ser Asn Asp Thr Val Pro Ile Lys Phe Lys Lys Cys Cys
                565                 570                 575

tat gga tat tgc att gat ctg ctg gaa aag ata gca gaa gac atg aac     1776
Tyr Gly Tyr Cys Ile Asp Leu Leu Glu Lys Ile Ala Glu Asp Met Asn
            580                 585                 590

ttt gac ttc gac ctc tat att gta ggg gat gga aag tat gga gca tgg     1824
Phe Asp Phe Asp Leu Tyr Ile Val Gly Asp Gly Lys Tyr Gly Ala Trp
        595                 600                 605

aaa aat ggg cac tgg act ggg cta gtg ggt gat ctc ctg aga ggg act     1872
Lys Asn Gly His Trp Thr Gly Leu Val Gly Asp Leu Leu Arg Gly Thr
    610                 615                 620

gcc cac atg gca gtc act tcc ttt agc atc aat act gca cgg agc cag     1920
Ala His Met Ala Val Thr Ser Phe Ser Ile Asn Thr Ala Arg Ser Gln
625                 630                 635                 640

gtg ata gac ttc acc agc cct ttc ttc tcc acc agc ttg ggc atc tta     1968
Val Ile Asp Phe Thr Ser Pro Phe Phe Ser Thr Ser Leu Gly Ile Leu
                645                 650                 655

gtg agg acc cga gat aca gca gct ccc att gga gcc ttc atg tgg cca     2016
Val Arg Thr Arg Asp Thr Ala Ala Pro Ile Gly Ala Phe Met Trp Pro
            660                 665                 670

ctc cac tgg aca atg tgg ctg ggg att ttt gtg gct ctg cac atc act     2064
Leu His Trp Thr Met Trp Leu Gly Ile Phe Val Ala Leu His Ile Thr
        675                 680                 685

gcc gtc ttc ctc act ctg tat gaa tgg aag agt cca ttt ggt ttg act     2112
Ala Val Phe Leu Thr Leu Tyr Glu Trp Lys Ser Pro Phe Gly Leu Thr
    690                 695                 700

ccc aag ggg cga aat aga agt aaa gtc ttc tcc ttt tct tca gcc ttg     2160
Pro Lys Gly Arg Asn Arg Ser Lys Val Phe Ser Phe Ser Ser Ala Leu
705                 710                 715                 720

aac atc tgt tat gcc ctc ttg ttt ggc aga aca gtg gcc atc aaa cct     2208
Asn Ile Cys Tyr Ala Leu Leu Phe Gly Arg Thr Val Ala Ile Lys Pro
                725                 730                 735

cca aaa tgt tgg act gga agg ttt cta atg aac ctt tgg gcc att ttc     2256
Pro Lys Cys Trp Thr Gly Arg Phe Leu Met Asn Leu Trp Ala Ile Phe
            740                 745                 750
```

-continued

```
tgt atg ttt tgc ctt tcc aca tac acg gca aac ttg gct gct gtc atg     2304
Cys Met Phe Cys Leu Ser Thr Tyr Thr Ala Asn Leu Ala Ala Val Met
        755                 760                 765

gta ggt gag aag atc tat gaa gag ctt tct gga ata cat gac ccc aag     2352
Val Gly Glu Lys Ile Tyr Glu Glu Leu Ser Gly Ile His Asp Pro Lys
    770                 775                 780

tta cat cat cct tcc caa gga ttc cgc ttt gga act gtc cga gaa agc     2400
Leu His His Pro Ser Gln Gly Phe Arg Phe Gly Thr Val Arg Glu Ser
785                 790                 795                 800

agt gct gaa gat tat gtg aga caa agt ttc cca gag atg cat gaa tat     2448
Ser Ala Glu Asp Tyr Val Arg Gln Ser Phe Pro Glu Met His Glu Tyr
                805                 810                 815

atg aga agg tac aat gtt cca gcc acc cct gat gga gtg gag tat ctg     2496
Met Arg Arg Tyr Asn Val Pro Ala Thr Pro Asp Gly Val Glu Tyr Leu
            820                 825                 830

aag aat gat cca gag aaa cta gac gcc ttc atc atg gac aaa gcc ctt     2544
Lys Asn Asp Pro Glu Lys Leu Asp Ala Phe Ile Met Asp Lys Ala Leu
        835                 840                 845

ctg gat tat gaa gtg tca ata gat gct gac tgc aaa ctt ctc act gtg     2592
Leu Asp Tyr Glu Val Ser Ile Asp Ala Asp Cys Lys Leu Leu Thr Val
    850                 855                 860

ggg aag cca ttt gcc ata gaa gga tac ggc att ggc ctc cca ccc aac     2640
Gly Lys Pro Phe Ala Ile Glu Gly Tyr Gly Ile Gly Leu Pro Pro Asn
865                 870                 875                 880

tct cca ttg acc gcc aac ata tcc gag cta atc agt caa tac aag tca     2688
Ser Pro Leu Thr Ala Asn Ile Ser Glu Leu Ile Ser Gln Tyr Lys Ser
                885                 890                 895

cat ggg ttt atg gat atg ctc cat gac aag tgg tac agg gtg gtt ccc     2736
His Gly Phe Met Asp Met Leu His Asp Lys Trp Tyr Arg Val Val Pro
            900                 905                 910

tgt ggc aag aga agt ttt gct gtc acg gag act ttg caa atg ggc atc     2784
Cys Gly Lys Arg Ser Phe Ala Val Thr Glu Thr Leu Gln Met Gly Ile
        915                 920                 925

aaa cac ttc tct ggg ctc ttt gtg ctg ctg tgc att gga ttt ggt ctg     2832
Lys His Phe Ser Gly Leu Phe Val Leu Leu Cys Ile Gly Phe Gly Leu
    930                 935                 940

tcc att ttg acc acc att ggt gag cac ata gta tac agg ctg ctg cta     2880
Ser Ile Leu Thr Thr Ile Gly Glu His Ile Val Tyr Arg Leu Leu Leu
945                 950                 955                 960

cca cga atc aaa aac aaa tcc aag ctg caa tac tgg ctc cac acc agc     2928
Pro Arg Ile Lys Asn Lys Ser Lys Leu Gln Tyr Trp Leu His Thr Ser
                965                 970                 975

cag aga tta cac aga gca ata aat aca tca ttt ata gag gaa aag cag     2976
Gln Arg Leu His Arg Ala Ile Asn Thr Ser Phe Ile Glu Glu Lys Gln
            980                 985                 990

cag cat ttc aag acc aaa cgt gtg gaa aag agg tct aat gtg gga ccc     3024
Gln His Phe Lys Thr Lys Arg Val Glu Lys Arg Ser Asn Val Gly Pro
        995                 1000                1005

cgt cag ctt acc gta tgg aat act tcc aat ctg agt cat gac aac cga     3072
Arg Gln Leu Thr Val Trp Asn Thr Ser Asn Leu Ser His Asp Asn Arg
    1010                1015                1020

cgg aaa tac atc ttt agt gat gag gaa gga caa aac cag ctg ggc atc     3120
Arg Lys Tyr Ile Phe Ser Asp Glu Glu Gly Gln Asn Gln Leu Gly Ile
1025                1030                1035                1040

cag atc cac cag gac atc ccc ctc cct cca agg aga aga gag ctc cct     3168
Gln Ile His Gln Asp Ile Pro Leu Pro Pro Arg Arg Arg Glu Leu Pro
                1045                1050                1055
```

-continued

```
gcc ttg cgg acc acc aat ggg aaa gca gac tcc cta aat gta tct cgg     3216
Ala Leu Arg Thr Thr Asn Gly Lys Ala Asp Ser Leu Asn Val Ser Arg
            1060                1065                1070

aac tca gtg atg cag gaa ctc tca gag ctc gag aag cag att cag gtg     3264
Asn Ser Val Met Gln Glu Leu Ser Glu Leu Glu Lys Gln Ile Gln Val
        1075                1080                1085

atc cgt cag gag ctg cag ctg gct gtg agc agg aaa acg gag ctg gag     3312
Ile Arg Gln Glu Leu Gln Leu Ala Val Ser Arg Lys Thr Glu Leu Glu
    1090                1095                1100

gag tat caa agg aca agt cgg act tgt gag tcc tag gtgaccacac          3358
Glu Tyr Gln Arg Thr Ser Arg Thr Cys Glu Ser  *
1105                1110                1115

tgcttccctt tctcagttcc tgaccttcct ctgagccctt gagacacttt gtaatgctct   3418

tttgtaacta tcgacaaagg tgtggggaag ctgaggtcta ggtcttctta aaggtcaagt   3478

ctgctctccc tcgcctaaag tgcagcagca gctcctctca agctcactct ctaggtctcc   3538

agggtaggag tgtttttcta gcaagaatct tagtcaggag taagctctgt gcgagagatc   3598

tgtgaataac cagataaccc cagctgccgt taaccttttc accaggtgcc acagtaatat   3658

ttctggtttt tagccctttc tctgcactac caacaagaga taaaattgtt actcacactt   3718

atgtcttact gggttgctgg ttttcatcgt aacacagaac gaggttatct agggttgtag   3778

cttttgatac aactccccga tctagattta ttcctacatt ctgaatgggg agcaggtaag   3838

agcagagcac ctcccactgg gggtggggta tttaaaaatt aactcattag tatcataaac   3898

gtcaaggatt gattggacca ggcaagagcc atgtttttga gaaggttctg gatctctgac   3958

tccatcctga ctgtttagta agagcatgct tacaccctac tgtgaaaagg ggaggggatg   4018

tggtaagcag aaacagaaga caggcagcag aggcattaaa aatgcatacc atgctttcag   4078

aacaaaagct ctgggccaga aaggcaattt ggctaaaaaa tgaataagac tacttctaat   4138

gtaactaagc atctccacta tggtgtgtgc cttttataaa ggaaaaaaaa aaaaaaagg    4197
```

<210> SEQ ID NO 57
<211> LENGTH: 1115
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 57

```
Met Arg Arg Leu Ser Leu Trp Trp Leu Leu Ser Arg Val Cys Leu Leu
 1               5                  10                  15

Leu Pro Pro Pro Cys Ala Leu Val Leu Ala Gly Val Pro Ser Ser Ser
            20                  25                  30

Ser His Pro Gln Pro Cys Gln Ile Leu Lys Arg Ile Gly His Ala Val
        35                  40                  45

Arg Val Gly Ala Val His Leu Gln Pro Trp Thr Thr Ala Pro Arg Ala
    50                  55                  60

Ala Ser Arg Ala Pro Asp Asp Ser Arg Ala Gly Ala Gln Arg Asp Glu
65                  70                  75                  80

Pro Glu Pro Gly Thr Arg Arg Ser Pro Ala Pro Ser Pro Gly Ala Arg
                85                  90                  95

Trp Leu Gly Ser Thr Leu His Gly Arg Gly Pro Pro Gly Ser Arg Lys
            100                 105                 110

Pro Gly Glu Gly Ala Arg Ala Glu Ala Leu Trp Pro Arg Asp Ala Leu
        115                 120                 125
```

-continued

```
Leu Phe Ala Val Asp Asn Leu Asn Arg Val Glu Gly Leu Leu Pro Tyr
    130                 135                 140

Asn Leu Ser Leu Glu Val Val Met Ala Ile Glu Ala Gly Leu Gly Asp
145                 150                 155                 160

Leu Pro Leu Leu Pro Phe Ser Ser Pro Ser Ser Pro Trp Ser Ser Asp
                165                 170                 175

Pro Phe Ser Phe Leu Gln Ser Val Cys His Thr Val Val Val Gln Gly
            180                 185                 190

Val Ser Ala Leu Leu Ala Phe Pro Gln Ser Gln Gly Glu Met Met Glu
        195                 200                 205

Leu Asp Leu Val Ser Leu Val Leu His Ile Pro Val Ile Ser Ile Val
    210                 215                 220

Arg His Glu Phe Pro Arg Glu Ser Gln Asn Pro Leu His Leu Gln Leu
225                 230                 235                 240

Ser Leu Glu Asn Ser Leu Ser Ser Asp Ala Asp Val Thr Val Ser Ile
                245                 250                 255

Leu Thr Met Asn Asn Trp Tyr Asn Phe Ser Leu Leu Leu Cys Gln Glu
            260                 265                 270

Asp Trp Asn Ile Thr Asp Phe Leu Leu Leu Thr Gln Asn Asn Ser Lys
        275                 280                 285

Phe His Leu Gly Ser Ile Ile Asn Ile Thr Ala Asn Leu Pro Ser Thr
    290                 295                 300

Gln Asp Leu Leu Ser Phe Leu Gln Ile Gln Leu Glu Ser Ile Lys Asn
305                 310                 315                 320

Ser Thr Pro Thr Val Val Met Phe Gly Cys Asp Met Glu Ser Ile Arg
                325                 330                 335

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

Arg Trp Val Leu Gly Asp Ser Gln Asn Met Glu Glu Leu Arg Thr Glu
        355                 360                 365

Gly Leu Pro Leu Gly Leu Ile Ala His Gly Lys Thr Thr Gln Ser Val
    370                 375                 380

Phe Glu His Tyr Val Gln Asp Ala Met Glu Leu Val Ala Arg Ala Val
385                 390                 395                 400

Ala Thr Ala Thr Met Ile Gln Pro Glu Leu Ala Leu Ile Pro Ser Thr
                405                 410                 415

Met Asn Cys Met Glu Val Glu Thr Thr Asn Leu Thr Ser Gly Gln Tyr
            420                 425                 430

Leu Ser Arg Phe Leu Ala Asn Thr Thr Phe Arg Gly Leu Ser Gly Ser
        435                 440                 445

Ile Arg Val Lys Gly Ser Thr Ile Val Ser Ser Glu Asn Asn Phe Phe
    450                 455                 460

Ile Trp Asn Leu Gln His Asp Pro Met Gly Lys Pro Met Trp Thr Arg
465                 470                 475                 480

Leu Gly Ser Trp Gln Gly Arg Lys Ile Val Met Asp Tyr Gly Ile Trp
                485                 490                 495

Pro Glu Gln Ala Gln Arg His Lys Thr His Phe Gln His Pro Ser Lys
            500                 505                 510

Leu His Leu Arg Val Val Thr Leu Ile Glu His Pro Phe Val Phe Thr
        515                 520                 525

Arg Glu Val Asp Asp Glu Gly Leu Cys Pro Ala Gly Gln Leu Cys Leu
```

-continued

```
    530                 535                 540

Asp Pro Met Thr Asn Asp Ser Ser Thr Leu Asp Ser Leu Phe Ser Ser
545                 550                 555                 560

Leu His Ser Ser Asn Asp Thr Val Pro Ile Lys Phe Lys Lys Cys Cys
                565                 570                 575

Tyr Gly Tyr Cys Ile Asp Leu Leu Glu Lys Ile Ala Glu Asp Met Asn
            580                 585                 590

Phe Asp Phe Asp Leu Tyr Ile Val Gly Asp Gly Lys Tyr Gly Ala Trp
        595                 600                 605

Lys Asn Gly His Trp Thr Gly Leu Val Gly Asp Leu Leu Arg Gly Thr
    610                 615                 620

Ala His Met Ala Val Thr Ser Phe Ser Ile Asn Thr Ala Arg Ser Gln
625                 630                 635                 640

Val Ile Asp Phe Thr Ser Pro Phe Phe Ser Thr Ser Leu Gly Ile Leu
                645                 650                 655

Val Arg Thr Arg Asp Thr Ala Ala Pro Ile Gly Ala Phe Met Trp Pro
            660                 665                 670

Leu His Trp Thr Met Trp Leu Gly Ile Phe Val Ala Leu His Ile Thr
        675                 680                 685

Ala Val Phe Leu Thr Leu Tyr Glu Trp Lys Ser Pro Phe Gly Leu Thr
    690                 695                 700

Pro Lys Gly Arg Asn Arg Ser Lys Val Phe Ser Phe Ser Ser Ala Leu
705                 710                 715                 720

Asn Ile Cys Tyr Ala Leu Leu Phe Gly Arg Thr Val Ala Ile Lys Pro
                725                 730                 735

Pro Lys Cys Trp Thr Gly Arg Phe Leu Met Asn Leu Trp Ala Ile Phe
            740                 745                 750

Cys Met Phe Cys Leu Ser Thr Tyr Thr Ala Asn Leu Ala Ala Val Met
        755                 760                 765

Val Gly Glu Lys Ile Tyr Glu Glu Leu Ser Gly Ile His Asp Pro Lys
    770                 775                 780

Leu His His Pro Ser Gln Gly Phe Arg Phe Gly Thr Val Arg Glu Ser
785                 790                 795                 800

Ser Ala Glu Asp Tyr Val Arg Gln Ser Phe Pro Glu Met His Glu Tyr
                805                 810                 815

Met Arg Arg Tyr Asn Val Pro Ala Thr Pro Asp Gly Val Glu Tyr Leu
            820                 825                 830

Lys Asn Asp Pro Glu Lys Leu Asp Ala Phe Ile Met Asp Lys Ala Leu
        835                 840                 845

Leu Asp Tyr Glu Val Ser Ile Asp Ala Asp Cys Lys Leu Leu Thr Val
    850                 855                 860

Gly Lys Pro Phe Ala Ile Glu Gly Tyr Gly Ile Gly Leu Pro Pro Asn
865                 870                 875                 880

Ser Pro Leu Thr Ala Asn Ile Ser Glu Leu Ile Ser Gln Tyr Lys Ser
                885                 890                 895

His Gly Phe Met Asp Met Leu His Asp Lys Trp Tyr Arg Val Val Pro
            900                 905                 910

Cys Gly Lys Arg Ser Phe Ala Val Thr Glu Thr Leu Gln Met Gly Ile
        915                 920                 925

Lys His Phe Ser Gly Leu Phe Val Leu Leu Cys Ile Gly Phe Gly Leu
    930                 935                 940
```

-continued

```
Ser Ile Leu Thr Thr Ile Gly Glu His Ile Val Tyr Arg Leu Leu Leu
945                 950                 955                 960

Pro Arg Ile Lys Asn Lys Ser Lys Leu Gln Tyr Trp Leu His Thr Ser
                965                 970                 975

Gln Arg Leu His Arg Ala Ile Asn Thr Ser Phe Ile Glu Glu Lys Gln
            980                 985                 990

Gln His Phe Lys Thr Lys Arg Val Glu Lys Arg Ser Asn Val Gly Pro
        995                 1000                1005

Arg Gln Leu Thr Val Trp Asn Thr Ser Asn Leu Ser His Asp Asn Arg
    1010                1015                1020

Arg Lys Tyr Ile Phe Ser Asp Glu Glu Gly Gln Asn Gln Leu Gly Ile
1025                1030                1035                1040

Gln Ile His Gln Asp Ile Pro Leu Pro Pro Arg Arg Arg Glu Leu Pro
                1045                1050                1055

Ala Leu Arg Thr Thr Asn Gly Lys Ala Asp Ser Leu Asn Val Ser Arg
            1060                1065                1070

Asn Ser Val Met Gln Glu Leu Ser Glu Leu Glu Lys Gln Ile Gln Val
        1075                1080                1085

Ile Arg Gln Glu Leu Gln Leu Ala Val Ser Arg Lys Thr Glu Leu Glu
    1090                1095                1100

Glu Tyr Gln Arg Thr Ser Arg Thr Cys Glu Ser
1105                1110                1115


<210> SEQ ID NO 58
<211> LENGTH: 3348
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(3348)

<400> SEQUENCE: 58

atg agg aga ctg agt ttg tgg tgg ctg ctg agc agg gtc tgt ctg ctg       48
Met Arg Arg Leu Ser Leu Trp Trp Leu Leu Ser Arg Val Cys Leu Leu
 1               5                  10                  15

ttg ccg ccg ccc tgc gca ctg gtg ctg gcc ggg gtg ccc agc tcc tcc       96
Leu Pro Pro Pro Cys Ala Leu Val Leu Ala Gly Val Pro Ser Ser Ser
            20                  25                  30

tcg cac ccg cag ccc tgc cag atc ctc aag cgc atc ggg cac gcg gtg      144
Ser His Pro Gln Pro Cys Gln Ile Leu Lys Arg Ile Gly His Ala Val
        35                  40                  45

agg gtg ggc gcg gtg cac ttg cag ccc tgg acc acc gcc ccc cgc gcg      192
Arg Val Gly Ala Val His Leu Gln Pro Trp Thr Thr Ala Pro Arg Ala
    50                  55                  60

gcc agc cgc gct ccg gac gac agc cga gca gga gcc cag agg gat gag      240
Ala Ser Arg Ala Pro Asp Asp Ser Arg Ala Gly Ala Gln Arg Asp Glu
65                  70                  75                  80

ccg gag cca ggg act agg cgg tcc ccg gcg ccc tcg ccg ggc gca cgc      288
Pro Glu Pro Gly Thr Arg Arg Ser Pro Ala Pro Ser Pro Gly Ala Arg
                85                  90                  95

tgg ttg ggg agc acc ctg cat ggc cgg ggg ccg ccg ggc tcc cgt aag      336
Trp Leu Gly Ser Thr Leu His Gly Arg Gly Pro Pro Gly Ser Arg Lys
            100                 105                 110

ccc ggg gag ggc gcc agg gcg gag gcc ctg tgg cca cgg gac gcc ctc      384
Pro Gly Glu Gly Ala Arg Ala Glu Ala Leu Trp Pro Arg Asp Ala Leu
        115                 120                 125
```

-continued

```
cta ttt gcc gtg gac aac ctg aac cgc gtg gaa ggg ctg cta ccc tac      432
Leu Phe Ala Val Asp Asn Leu Asn Arg Val Glu Gly Leu Leu Pro Tyr
    130                 135                 140

aac ctg tct ttg gaa gta gtg atg gcc atc gag gca ggc ctg ggc gat      480
Asn Leu Ser Leu Glu Val Val Met Ala Ile Glu Ala Gly Leu Gly Asp
145                 150                 155                 160

ctg cca ctt ttg ccc ttc tcc tcc cct agt tcg cca tgg agc agt gac      528
Leu Pro Leu Leu Pro Phe Ser Ser Pro Ser Ser Pro Trp Ser Ser Asp
                165                 170                 175

cct ttc tcc ttc ctg caa agt gtg tgc cat acc gtg gtg gtg caa ggg      576
Pro Phe Ser Phe Leu Gln Ser Val Cys His Thr Val Val Val Gln Gly
            180                 185                 190

gtg tcg gcg ctg ctc gcc ttc ccc cag agc cag ggc gaa atg atg gag      624
Val Ser Ala Leu Leu Ala Phe Pro Gln Ser Gln Gly Glu Met Met Glu
        195                 200                 205

ctc gac ttg gtc agc tta gtc ctg cac att cca gtg atc agc atc gtg      672
Leu Asp Leu Val Ser Leu Val Leu His Ile Pro Val Ile Ser Ile Val
    210                 215                 220

cgc cac gag ttt ccg cgg gag agt cag aat ccc ctt cac cta caa ctg      720
Arg His Glu Phe Pro Arg Glu Ser Gln Asn Pro Leu His Leu Gln Leu
225                 230                 235                 240

agt tta gaa aat tca tta agt tct gat gct gat gtc act gtc tca atc      768
Ser Leu Glu Asn Ser Leu Ser Ser Asp Ala Asp Val Thr Val Ser Ile
                245                 250                 255

ctg acc atg aac aac tgg tac aat ttt agc ttg ttg ctg tgc cag gaa      816
Leu Thr Met Asn Asn Trp Tyr Asn Phe Ser Leu Leu Leu Cys Gln Glu
            260                 265                 270

gac tgg aac atc acc gac ttc ctc ctc ctt acc cag aat aat tcc aag      864
Asp Trp Asn Ile Thr Asp Phe Leu Leu Leu Thr Gln Asn Asn Ser Lys
        275                 280                 285

ttc cac ctt ggt tct atc atc aac atc acc gct aac ctc ccc tcc acc      912
Phe His Leu Gly Ser Ile Ile Asn Ile Thr Ala Asn Leu Pro Ser Thr
    290                 295                 300

cag gac ctc ttg agc ttc cta cag atc cag ctt gag agt att aag aac      960
Gln Asp Leu Leu Ser Phe Leu Gln Ile Gln Leu Glu Ser Ile Lys Asn
305                 310                 315                 320

agc aca ccc aca gtg gtg atg ttt ggc tgc gac atg gaa agt atc cgg     1008
Ser Thr Pro Thr Val Val Met Phe Gly Cys Asp Met Glu Ser Ile Arg
                325                 330                 335

cgg att ttc gaa att aca acc cag ttt ggg gtc atg ccc cct gaa ctt     1056
Arg Ile Phe Glu Ile Thr Thr Gln Phe Gly Val Met Pro Pro Glu Leu
            340                 345                 350

cgt tgg gtg ctg gga gat tcc cag aat atg gag gaa ctg agg aca gag     1104
Arg Trp Val Leu Gly Asp Ser Gln Asn Met Glu Glu Leu Arg Thr Glu
        355                 360                 365

ggt ctg ccc tta gga ctc att gct cat gga aaa aca aca cag tct gtc     1152
Gly Leu Pro Leu Gly Leu Ile Ala His Gly Lys Thr Thr Gln Ser Val
    370                 375                 380

ttt gag cac tac gta caa gat gct atg gag ctg gtc gca aga gct gta     1200
Phe Glu His Tyr Val Gln Asp Ala Met Glu Leu Val Ala Arg Ala Val
385                 390                 395                 400

gcc aca gcc acc atg atc caa cca gaa ctt gct ctc att ccc agc acg     1248
Ala Thr Ala Thr Met Ile Gln Pro Glu Leu Ala Leu Ile Pro Ser Thr
                405                 410                 415

atg aac tgc atg gag gtg gaa act aca aat ctc act tca gga caa tat     1296
Met Asn Cys Met Glu Val Glu Thr Thr Asn Leu Thr Ser Gly Gln Tyr
            420                 425                 430
```

-continued

```
tta tca agg ttt cta gcc aat acc act ttc aga ggc ctc agt ggt tcc     1344
Leu Ser Arg Phe Leu Ala Asn Thr Thr Phe Arg Gly Leu Ser Gly Ser
        435                 440                 445

att aga gta aaa ggt tcc acc atc gtc agc tca gaa aac aac ttt ttc     1392
Ile Arg Val Lys Gly Ser Thr Ile Val Ser Ser Glu Asn Asn Phe Phe
    450                 455                 460

atc tgg aat ctt caa cat gac ccc atg gga aag cca atg tgg acc cgc     1440
Ile Trp Asn Leu Gln His Asp Pro Met Gly Lys Pro Met Trp Thr Arg
465                 470                 475                 480

ttg ggc agc tgg cag ggg aga aag att gtc atg gac tat gga ata tgg     1488
Leu Gly Ser Trp Gln Gly Arg Lys Ile Val Met Asp Tyr Gly Ile Trp
                485                 490                 495

cca gag cag gcc cag aga cac aaa acc cac ttc caa cat cca agt aag     1536
Pro Glu Gln Ala Gln Arg His Lys Thr His Phe Gln His Pro Ser Lys
            500                 505                 510

cta cac ttg aga gtg gtt acc ctg att gag cat cct ttt gtc ttc aca     1584
Leu His Leu Arg Val Val Thr Leu Ile Glu His Pro Phe Val Phe Thr
        515                 520                 525

agg gag gta gat gat gaa ggc ttg tgc cct gct ggc caa ctc tgt cta     1632
Arg Glu Val Asp Asp Glu Gly Leu Cys Pro Ala Gly Gln Leu Cys Leu
    530                 535                 540

gac ccc atg act aat gac tct tcc aca ctg gac agc ctt ttt agc agc     1680
Asp Pro Met Thr Asn Asp Ser Ser Thr Leu Asp Ser Leu Phe Ser Ser
545                 550                 555                 560

ctc cat agc agt aat gat aca gtg ccc att aaa ttc aag aag tgc tgc     1728
Leu His Ser Ser Asn Asp Thr Val Pro Ile Lys Phe Lys Lys Cys Cys
                565                 570                 575

tat gga tat tgc att gat ctg ctg gaa aag ata gca gaa gac atg aac     1776
Tyr Gly Tyr Cys Ile Asp Leu Leu Glu Lys Ile Ala Glu Asp Met Asn
            580                 585                 590

ttt gac ttc gac ctc tat att gta ggg gat gga aag tat gga gca tgg     1824
Phe Asp Phe Asp Leu Tyr Ile Val Gly Asp Gly Lys Tyr Gly Ala Trp
        595                 600                 605

aaa aat ggg cac tgg act ggg cta gtg ggt gat ctc ctg aga ggg act     1872
Lys Asn Gly His Trp Thr Gly Leu Val Gly Asp Leu Leu Arg Gly Thr
    610                 615                 620

gcc cac atg gca gtc act tcc ttt agc atc aat act gca cgg agc cag     1920
Ala His Met Ala Val Thr Ser Phe Ser Ile Asn Thr Ala Arg Ser Gln
625                 630                 635                 640

gtg ata gac ttc acc agc cct ttc ttc tcc acc agc ttg ggc atc tta     1968
Val Ile Asp Phe Thr Ser Pro Phe Phe Ser Thr Ser Leu Gly Ile Leu
                645                 650                 655

gtg agg acc cga gat aca gca gct ccc att gga gcc ttc atg tgg cca     2016
Val Arg Thr Arg Asp Thr Ala Ala Pro Ile Gly Ala Phe Met Trp Pro
            660                 665                 670

ctc cac tgg aca atg tgg ctg ggg att ttt gtg gct ctg cac atc act     2064
Leu His Trp Thr Met Trp Leu Gly Ile Phe Val Ala Leu His Ile Thr
        675                 680                 685

gcc gtc ttc ctc act ctg tat gaa tgg aag agt cca ttt ggt ttg act     2112
Ala Val Phe Leu Thr Leu Tyr Glu Trp Lys Ser Pro Phe Gly Leu Thr
    690                 695                 700

ccc aag ggg cga aat aga agt aaa gtc ttc tcc ttt tct tca gcc ttg     2160
Pro Lys Gly Arg Asn Arg Ser Lys Val Phe Ser Phe Ser Ser Ala Leu
705                 710                 715                 720

aac atc tgt tat gcc ctc ttg ttt ggc aga aca gtg gcc atc aaa cct     2208
Asn Ile Cys Tyr Ala Leu Leu Phe Gly Arg Thr Val Ala Ile Lys Pro
                725                 730                 735
```

-continued

```
cca aaa tgt tgg act gga agg ttt cta atg aac ctt tgg gcc att ttc     2256
Pro Lys Cys Trp Thr Gly Arg Phe Leu Met Asn Leu Trp Ala Ile Phe
            740                 745                 750

tgt atg ttt tgc ctt tcc aca tac acg gca aac ttg gct gct gtc atg     2304
Cys Met Phe Cys Leu Ser Thr Tyr Thr Ala Asn Leu Ala Ala Val Met
        755                 760                 765

gta ggt gag aag atc tat gaa gag ctt tct gga ata cat gac ccc aag     2352
Val Gly Glu Lys Ile Tyr Glu Glu Leu Ser Gly Ile His Asp Pro Lys
    770                 775                 780

tta cat cat cct tcc caa gga ttc cgc ttt gga act gtc cga gaa agc     2400
Leu His His Pro Ser Gln Gly Phe Arg Phe Gly Thr Val Arg Glu Ser
785                 790                 795                 800

agt gct gaa gat tat gtg aga caa agt ttc cca gag atg cat gaa tat     2448
Ser Ala Glu Asp Tyr Val Arg Gln Ser Phe Pro Glu Met His Glu Tyr
                805                 810                 815

atg aga agg tac aat gtt cca gcc acc cct gat gga gtg gag tat ctg     2496
Met Arg Arg Tyr Asn Val Pro Ala Thr Pro Asp Gly Val Glu Tyr Leu
            820                 825                 830

aag aat gat cca gag aaa cta gac gcc ttc atc atg gac aaa gcc ctt     2544
Lys Asn Asp Pro Glu Lys Leu Asp Ala Phe Ile Met Asp Lys Ala Leu
        835                 840                 845

ctg gat tat gaa gtg tca ata gat gct gac tgc aaa ctt ctc act gtg     2592
Leu Asp Tyr Glu Val Ser Ile Asp Ala Asp Cys Lys Leu Leu Thr Val
    850                 855                 860

ggg aag cca ttt gcc ata gaa gga tac ggc att ggc ctc cca ccc aac     2640
Gly Lys Pro Phe Ala Ile Glu Gly Tyr Gly Ile Gly Leu Pro Pro Asn
865                 870                 875                 880

tct cca ttg acc gcc aac ata tcc gag cta atc agt caa tac aag tca     2688
Ser Pro Leu Thr Ala Asn Ile Ser Glu Leu Ile Ser Gln Tyr Lys Ser
                885                 890                 895

cat ggg ttt atg gat atg ctc cat gac aag tgg tac agg gtg gtt ccc     2736
His Gly Phe Met Asp Met Leu His Asp Lys Trp Tyr Arg Val Val Pro
            900                 905                 910

tgt ggc aag aga agt ttt gct gtc acg gag act ttg caa atg ggc atc     2784
Cys Gly Lys Arg Ser Phe Ala Val Thr Glu Thr Leu Gln Met Gly Ile
        915                 920                 925

aaa cac ttc tct ggg ctc ttt gtg ctg ctg tgc att gga ttt ggt ctg     2832
Lys His Phe Ser Gly Leu Phe Val Leu Leu Cys Ile Gly Phe Gly Leu
    930                 935                 940

tcc att ttg acc acc att ggt gag cac ata gta tac agg ctg ctg cta     2880
Ser Ile Leu Thr Thr Ile Gly Glu His Ile Val Tyr Arg Leu Leu Leu
945                 950                 955                 960

cca cga atc aaa aac aaa tcc aag ctg caa tac tgg ctc cac acc agc     2928
Pro Arg Ile Lys Asn Lys Ser Lys Leu Gln Tyr Trp Leu His Thr Ser
                965                 970                 975

cag aga tta cac aga gca ata aat aca tca ttt ata gag gaa aag cag     2976
Gln Arg Leu His Arg Ala Ile Asn Thr Ser Phe Ile Glu Glu Lys Gln
            980                 985                 990

cag cat ttc aag acc aaa cgt gtg gaa aag agg tct aat gtg gga ccc     3024
Gln His Phe Lys Thr Lys Arg Val Glu Lys Arg Ser Asn Val Gly Pro
        995                 1000                1005

cgt cag ctt acc gta tgg aat act tcc aat ctg agt cat gac aac cga     3072
Arg Gln Leu Thr Val Trp Asn Thr Ser Asn Leu Ser His Asp Asn Arg
    1010                1015                1020

cgg aaa tac atc ttt agt gat gag gaa gga caa aac cag ctg ggc atc     3120
Arg Lys Tyr Ile Phe Ser Asp Glu Glu Gly Gln Asn Gln Leu Gly Ile
1025                1030                1035                1040
```

-continued

```
cag atc cac cag gac atc ccc ctc cct cca agg aga aga gag ctc cct     3168
Gln Ile His Gln Asp Ile Pro Leu Pro Pro Arg Arg Arg Glu Leu Pro
                1045                1050                1055

gcc ttg cgg acc acc aat ggg aaa gca gac tcc cta aat gta tct cgg     3216
Ala Leu Arg Thr Thr Asn Gly Lys Ala Asp Ser Leu Asn Val Ser Arg
            1060                1065                1070

aac tca gtg atg cag gaa ctc tca gag ctc gag aag cag att cag gtg     3264
Asn Ser Val Met Gln Glu Leu Ser Glu Leu Glu Lys Gln Ile Gln Val
        1075                1080                1085

atc cgt cag gag ctg cag ctg gct gtg agc agg aaa acg gag ctg gag     3312
Ile Arg Gln Glu Leu Gln Leu Ala Val Ser Arg Lys Thr Glu Leu Glu
    1090                1095                1100

gag tat caa agg aca agt cgg act tgt gag tcc tag                     3348
Glu Tyr Gln Arg Thr Ser Arg Thr Cys Glu Ser  *
1105                1110                1115


<210> SEQ ID NO 59
<211> LENGTH: 1115
<212> TYPE: PRT
<213> ORGANISM: Rattus norvegicus

<400> SEQUENCE: 59

Met Arg Arg Leu Ser Leu Trp Trp Leu Leu Ser Arg Val Cys Leu Leu
 1               5                  10                  15

Leu Pro Pro Pro Cys Ala Leu Val Leu Ala Gly Val Pro Ser Ser Ser
            20                  25                  30

Ser His Pro Gln Pro Cys Gln Ile Leu Lys Arg Ile Gly His Ala Val
        35                  40                  45

Arg Val Gly Ala Val His Leu Gln Pro Trp Thr Thr Ala Pro Arg Ala
    50                  55                  60

Ala Ser Arg Ala Gln Glu Gly Gly Arg Ala Gly Ala Gln Arg Asp Asp
65                  70                  75                  80

Pro Glu Ser Gly Thr Trp Arg Pro Pro Ala Pro Ser Gln Gly Ala Arg
                85                  90                  95

Trp Leu Gly Ser Ala Leu His Gly Arg Gly Pro Pro Gly Ser Arg Lys
            100                 105                 110

Leu Gly Glu Gly Ala Gly Ala Glu Thr Leu Trp Pro Arg Asp Ala Leu
        115                 120                 125

Leu Phe Ala Val Glu Asn Leu Asn Arg Val Glu Gly Leu Leu Pro Tyr
    130                 135                 140

Asn Leu Ser Leu Glu Val Val Met Ala Ile Glu Ala Gly Leu Gly Asp
145                 150                 155                 160

Leu Pro Leu Met Pro Phe Ser Ser Pro Ser Ser Pro Trp Ser Ser Asp
                165                 170                 175

Pro Phe Ser Phe Leu Gln Ser Val Cys His Thr Val Val Val Gln Gly
            180                 185                 190

Val Ser Ala Leu Leu Ala Phe Pro Gln Ser Gln Gly Glu Met Met Glu
        195                 200                 205

Leu Asp Leu Val Ser Ser Val Leu His Ile Pro Val Leu Ser Ile Val
    210                 215                 220

Arg His Glu Phe Pro Arg Glu Ser Gln Asn Pro Leu His Leu Gln Leu
225                 230                 235                 240

Ser Leu Glu Asn Ser Leu Ser Ser Asp Ala Asp Val Thr Val Ser Ile
                245                 250                 255
```

-continued

```
Leu Thr Met Asn Asn Trp Tyr Asn Phe Ser Leu Leu Leu Cys Gln Glu
            260                 265                 270

Asp Trp Asn Ile Thr Asp Phe Leu Leu Leu Thr Glu Asn Asn Ser Lys
        275                 280                 285

Phe His Leu Glu Ser Val Ile Asn Ile Thr Ala Asn Leu Ser Ser Thr
    290                 295                 300

Lys Asp Leu Leu Ser Phe Leu Gln Val Gln Met Asp Asn Ile Arg Asn
305                 310                 315                 320

Ser Thr Pro Thr Met Val Met Phe Gly Cys Asp Met Asp Ser Ile Arg
                325                 330                 335

Gln Ile Phe Glu Met Ser Thr Gln Phe Gly Leu Ser Pro Pro Glu Leu
            340                 345                 350

His Trp Val Leu Gly Asp Ser Gln Asn Val Glu Glu Leu Arg Thr Glu
        355                 360                 365

Gly Leu Pro Leu Gly Leu Ile Ala His Gly Lys Thr Thr Gln Ser Val
    370                 375                 380

Phe Glu Tyr Tyr Val Gln Asp Ala Met Glu Leu Val Ala Arg Ala Val
385                 390                 395                 400

Ala Thr Ala Thr Met Ile Gln Pro Glu Leu Ala Leu Leu Pro Ser Thr
                405                 410                 415

Met Asn Cys Met Asp Val Lys Thr Thr Asn Leu Thr Ser Gly Gln Tyr
            420                 425                 430

Leu Ser Arg Phe Leu Ala Asn Thr Thr Phe Arg Gly Leu Ser Gly Ser
        435                 440                 445

Ile Lys Val Lys Gly Ser Thr Ile Ile Ser Ser Glu Asn Asn Phe Phe
    450                 455                 460

Ile Trp Asn Leu Gln His Asp Pro Met Gly Lys Pro Met Trp Thr Arg
465                 470                 475                 480

Leu Gly Ser Trp Gln Gly Gly Arg Ile Val Met Asp Ser Gly Ile Trp
                485                 490                 495

Pro Glu Gln Ala Gln Arg His Lys Thr His Phe Gln His Pro Asn Lys
            500                 505                 510

Leu His Leu Arg Val Val Thr Leu Ile Glu His Pro Phe Val Phe Thr
        515                 520                 525

Arg Glu Val Asp Asp Glu Gly Leu Cys Pro Ala Gly Gln Leu Cys Leu
    530                 535                 540

Asp Pro Met Thr Asn Asp Ser Ser Met Leu Asp Arg Leu Phe Ser Ser
545                 550                 555                 560

Leu His Ser Ser Asn Asp Thr Val Pro Ile Lys Phe Lys Lys Cys Cys
                565                 570                 575

Tyr Gly Tyr Cys Ile Asp Leu Leu Glu Gln Leu Ala Glu Asp Met Asn
            580                 585                 590

Phe Asp Phe Asp Leu Tyr Ile Val Gly Asp Gly Lys Tyr Gly Ala Trp
        595                 600                 605

Lys Asn Gly His Trp Thr Gly Leu Val Gly Asp Leu Leu Ser Gly Thr
    610                 615                 620

Ala Asn Met Ala Val Thr Ser Phe Ser Ile Asn Thr Ala Arg Ser Gln
625                 630                 635                 640

Val Ile Asp Phe Thr Ser Pro Phe Phe Ser Thr Ser Leu Gly Ile Leu
                645                 650                 655
```

-continued

```
Val Arg Thr Arg Asp Thr Ala Ala Pro Ile Gly Ala Phe Met Trp Pro
            660                 665                 670

Leu His Trp Thr Met Trp Leu Gly Ile Phe Val Ala Leu His Ile Thr
        675                 680                 685

Ala Ile Phe Leu Thr Leu Tyr Glu Trp Lys Ser Pro Phe Gly Met Thr
    690                 695                 700

Pro Lys Gly Arg Asn Arg Asn Lys Val Phe Ser Phe Ser Ser Ala Leu
705                 710                 715                 720

Asn Val Cys Tyr Ala Leu Leu Phe Gly Arg Thr Ala Ala Ile Lys Pro
                725                 730                 735

Pro Lys Cys Trp Thr Gly Arg Phe Leu Met Asn Leu Trp Ala Ile Phe
            740                 745                 750

Cys Met Phe Cys Leu Ser Thr Tyr Thr Ala Asn Leu Ala Ala Val Met
        755                 760                 765

Val Gly Glu Lys Ile Tyr Glu Glu Leu Ser Gly Ile His Asp Pro Lys
    770                 775                 780

Leu His His Pro Ser Gln Gly Phe Arg Phe Gly Thr Val Arg Glu Ser
785                 790                 795                 800

Ser Ala Glu Asp Tyr Val Arg Gln Ser Phe Pro Glu Met His Glu Tyr
                805                 810                 815

Met Arg Arg Tyr Asn Val Pro Ala Thr Pro Asp Gly Val Gln Tyr Leu
            820                 825                 830

Lys Asn Asp Pro Glu Lys Leu Asp Ala Phe Ile Met Asp Lys Ala Leu
        835                 840                 845

Leu Asp Tyr Glu Val Ser Ile Asp Ala Asp Cys Lys Leu Leu Thr Val
    850                 855                 860

Gly Lys Pro Phe Ala Ile Glu Gly Tyr Gly Ile Gly Leu Pro Pro Asn
865                 870                 875                 880

Ser Pro Leu Thr Ser Asn Ile Ser Glu Leu Ile Ser Gln Tyr Lys Ser
                885                 890                 895

His Gly Phe Met Asp Val Leu His Asp Lys Trp Tyr Lys Val Val Pro
            900                 905                 910

Cys Gly Lys Arg Ser Phe Ala Val Thr Glu Thr Leu Gln Met Gly Ile
        915                 920                 925

Lys His Phe Ser Gly Leu Phe Val Leu Leu Cys Ile Gly Phe Gly Leu
    930                 935                 940

Ser Ile Leu Thr Thr Ile Gly Glu His Ile Val His Arg Leu Leu Leu
945                 950                 955                 960

Pro Arg Ile Lys Asn Lys Ser Lys Leu Gln Tyr Trp Leu His Thr Ser
                965                 970                 975

Gln Arg Phe His Arg Ala Leu Asn Thr Ser Phe Val Glu Glu Lys Gln
            980                 985                 990

Pro Arg Ser Lys Thr Lys Arg Val Glu Lys Arg Ser Asn Leu Gly Pro
        995                 1000                1005

Gln Gln Leu Met Val Trp Asn Thr Ser Asn Leu Ser His Asp Asn Gln
    1010                1015                1020

Arg Lys Tyr Ile Phe Asn Asp Glu Glu Gly Gln Asn Gln Leu Gly Thr
1025                1030                1035                1040

Gln Ala His Gln Asp Ile Pro Leu Pro Gln Arg Arg Arg Glu Leu Pro
                1045                1050                1055

Ala Ser Leu Thr Thr Asn Gly Lys Ala Asp Ser Leu Asn Val Thr Arg
```

-continued

```
            1060                1065                1070

Ser Ser Val Ile Gln Glu Leu Ser Glu Leu Glu Lys Gln Ile Gln Val
        1075                1080                1085

Ile Arg Gln Glu Leu Gln Leu Ala Val Ser Arg Lys Thr Glu Leu Glu
    1090                1095                1100

Glu Tyr Gln Lys Thr Asn Arg Thr Cys Glu Ser
1105                1110                1115


<210> SEQ ID NO 60
<211> LENGTH: 2247
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (51)...(2048)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(2247)
<223> OTHER INFORMATION: n = A,T,C or G

<400> SEQUENCE: 60

gcacgaggca aaggacctcc agaccagagc cagccagcag caaaaagagc atg gag        56
Met Glu
 1

ctg agg agt aca gca gcc ccc aga gct gag ggc tac agc aac gtg ggc      104
Leu Arg Ser Thr Ala Ala Pro Arg Ala Glu Gly Tyr Ser Asn Val Gly
         5                  10                  15

ttc cag aat gaa gaa aac ttt ctt gag aac gag aac aca tca gga aac      152
Phe Gln Asn Glu Glu Asn Phe Leu Glu Asn Glu Asn Thr Ser Gly Asn
     20                  25                  30

aac tca ata aga agc aga gct gtg caa agc agg gag cac aca aac acc      200
Asn Ser Ile Arg Ser Arg Ala Val Gln Ser Arg Glu His Thr Asn Thr
 35                  40                  45                  50

aaa cag gat gaa gaa cag gtc aca gtt gag cag gat ttt cca aga aac      248
Lys Gln Asp Glu Glu Gln Val Thr Val Glu Gln Asp Phe Pro Arg Asn
                 55                  60                  65

aga gaa cac atg gag gat gat gat gag gag atg caa caa aaa ggg tgt      296
Arg Glu His Met Glu Asp Asp Asp Glu Glu Met Gln Gln Lys Gly Cys
             70                  75                  80

ttg gaa agg agg tat gac aca gta tgt ggt ttc tgt agg aaa cac aaa      344
Leu Glu Arg Arg Tyr Asp Thr Val Cys Gly Phe Cys Arg Lys His Lys
         85                  90                  95

aca act ctt cgg cac atc atc tgg ggc att tta tta gca ggt tat ctg      392
Thr Thr Leu Arg His Ile Ile Trp Gly Ile Leu Leu Ala Gly Tyr Leu
    100                 105                 110

gtt atg gtg att tcg gcc tgt gtg ctg aac ttt cac aga gcc ctt cct      440
Val Met Val Ile Ser Ala Cys Val Leu Asn Phe His Arg Ala Leu Pro
115                 120                 125                 130

ctt ttt gtg atc acc gtg gct gcc atc ttc ttt gtt gtc tgg gat cac      488
Leu Phe Val Ile Thr Val Ala Ala Ile Phe Phe Val Val Trp Asp His
                135                 140                 145

ctg atg gcc aaa tac gaa cat cga att gat gag atg ctg tct cct ggc      536
Leu Met Ala Lys Tyr Glu His Arg Ile Asp Glu Met Leu Ser Pro Gly
            150                 155                 160

aga agg ctt cta aac agc cat tgg ttc tgg ctg aag tgg gtg atc tgg      584
Arg Arg Leu Leu Asn Ser His Trp Phe Trp Leu Lys Trp Val Ile Trp
        165                 170                 175

agc tcc ctg gtc cta gca gtt att ttc tgg ttg gcc ttt gac act gcc      632
Ser Ser Leu Val Leu Ala Val Ile Phe Trp Leu Ala Phe Asp Thr Ala
```

-continued

```
    180                 185                 190

aaa ttg ggt caa cag cag ctg gtg tcc ttc ggt ggg ctc ata atg tac      680
Lys Leu Gly Gln Gln Gln Leu Val Ser Phe Gly Gly Leu Ile Met Tyr
195                 200                 205                 210

att gtc ctg tta ttt cta ttc tcc aag tac cca acc aga gtt tac tgg      728
Ile Val Leu Leu Phe Leu Phe Ser Lys Tyr Pro Thr Arg Val Tyr Trp
                215                 220                 225

aga cct gtc tta tgg gga atc ggg cta cag ttt ctt ctt ggg ctc ttg      776
Arg Pro Val Leu Trp Gly Ile Gly Leu Gln Phe Leu Leu Gly Leu Leu
            230                 235                 240

att cta agg act gac cct gga ttt ata gct ttt gat tgg ttg ggc aga      824
Ile Leu Arg Thr Asp Pro Gly Phe Ile Ala Phe Asp Trp Leu Gly Arg
        245                 250                 255

caa gtt cag gtc ctg ccg atc gtg gtt ttc ttc agc act gtg atg tcc      872
Gln Val Gln Val Leu Pro Ile Val Val Phe Phe Ser Thr Val Met Ser
    260                 265                 270

atg ctg tac tac ctg gga ctg atg cag tgg att att aga aag gtt gga      920
Met Leu Tyr Tyr Leu Gly Leu Met Gln Trp Ile Ile Arg Lys Val Gly
275                 280                 285                 290

tgg atc atg cta gtt act acg gga tca tct cct att gaa tct gta gtt      968
Trp Ile Met Leu Val Thr Thr Gly Ser Ser Pro Ile Glu Ser Val Val
                295                 300                 305

gct tct ggc aat ata ttt gtt gga caa acg gag tct cct ctg ctg gtc     1016
Ala Ser Gly Asn Ile Phe Val Gly Gln Thr Glu Ser Pro Leu Leu Val
            310                 315                 320

cga cca tat tta cct tac atc acc aag tct gaa ctc cac gcc atc atg     1064
Arg Pro Tyr Leu Pro Tyr Ile Thr Lys Ser Glu Leu His Ala Ile Met
        325                 330                 335

acc gcc ggg ttc tct acc att gct gga agc gtg cta ggt gca tac att     1112
Thr Ala Gly Phe Ser Thr Ile Ala Gly Ser Val Leu Gly Ala Tyr Ile
    340                 345                 350

tct ttt ggg gtt cca tcc tcc cac ttg tta aca gcg tca gtt atg tca     1160
Ser Phe Gly Val Pro Ser Ser His Leu Leu Thr Ala Ser Val Met Ser
355                 360                 365                 370

gca cct gcg tca ttg gct gct gct aaa ctc ttt tgg cct gag aca gaa     1208
Ala Pro Ala Ser Leu Ala Ala Ala Lys Leu Phe Trp Pro Glu Thr Glu
                375                 380                 385

aaa cct aaa ata acc ctc aag aat gcc atg aaa atg gaa agt ggt gat     1256
Lys Pro Lys Ile Thr Leu Lys Asn Ala Met Lys Met Glu Ser Gly Asp
            390                 395                 400

tca ggg aat ctt cta gaa gct gca aca cag gga gca tcc tcc tcc atc     1304
Ser Gly Asn Leu Leu Glu Ala Ala Thr Gln Gly Ala Ser Ser Ser Ile
        405                 410                 415

tcc ctg gtg gcc aac atc gct gtg aat ctg att gcc ttc ctg gcc ctg     1352
Ser Leu Val Ala Asn Ile Ala Val Asn Leu Ile Ala Phe Leu Ala Leu
    420                 425                 430

ctg tct ttt atg aat tca gcc ctg tcc tgg ttt gga aac atg ttt gac     1400
Leu Ser Phe Met Asn Ser Ala Leu Ser Trp Phe Gly Asn Met Phe Asp
435                 440                 445                 450

tac cca cag ctg agt ttt gag cta atc tgc tcc tac atc ttc atg ccc     1448
Tyr Pro Gln Leu Ser Phe Glu Leu Ile Cys Ser Tyr Ile Phe Met Pro
                455                 460                 465

ttt tcc ttc atg atg gga gtg gaa tgg cag gac agc ttt atg gtt gcc     1496
Phe Ser Phe Met Met Gly Val Glu Trp Gln Asp Ser Phe Met Val Ala
            470                 475                 480

aga ctc ata ggt tat aag acc ttc ttc aat gaa ttt gtg gct tat gag     1544
Arg Leu Ile Gly Tyr Lys Thr Phe Phe Asn Glu Phe Val Ala Tyr Glu
```

-continued

```
        485                 490                 495

cac ctc tca aaa tgg atc cac ttg agg aaa gaa ggt gga ccc aaa ttt     1592
His Leu Ser Lys Trp Ile His Leu Arg Lys Glu Gly Gly Pro Lys Phe
    500                 505                 510

gta aac ggt gtg cag caa tat ata tca att cgt tct gag ata atc gcc     1640
Val Asn Gly Val Gln Gln Tyr Ile Ser Ile Arg Ser Glu Ile Ile Ala
515                 520                 525                 530

act tac gct ctc tgt ggt ttt gcc aat atc ggg tcc cta gga atc gtg     1688
Thr Tyr Ala Leu Cys Gly Phe Ala Asn Ile Gly Ser Leu Gly Ile Val
                535                 540                 545

atc ggc gga ctc aca tcc atg gct cct tcc aga aag cgt gat atc gcc     1736
Ile Gly Gly Leu Thr Ser Met Ala Pro Ser Arg Lys Arg Asp Ile Ala
            550                 555                 560

tcg ggg gca gtg aga gct ctg att gcg ggg acc gtg gcc tgc ttc atg     1784
Ser Gly Ala Val Arg Ala Leu Ile Ala Gly Thr Val Ala Cys Phe Met
        565                 570                 575

aca gcc tgc atc gca ggc ata ctc tcc agc act cct gtg gac atc aac     1832
Thr Ala Cys Ile Ala Gly Ile Leu Ser Ser Thr Pro Val Asp Ile Asn
    580                 585                 590

tgc cat cac gtt tta gag aat gcc ttc aac tcc act ttc cct gga aac     1880
Cys His His Val Leu Glu Asn Ala Phe Asn Ser Thr Phe Pro Gly Asn
595                 600                 605                 610

aca acc aag gtg ata gct tgt tgc caa agt ctg ttg agc agc act gtt     1928
Thr Thr Lys Val Ile Ala Cys Cys Gln Ser Leu Leu Ser Ser Thr Val
                615                 620                 625

gcc aag ggt cct ggt gaa gtc atc cca gga gga aac cac agt ctg tat     1976
Ala Lys Gly Pro Gly Glu Val Ile Pro Gly Gly Asn His Ser Leu Tyr
            630                 635                 640

tct ttg aag ggc tgc tgc aca ttg ttg aat cca tcg acc ttt aac tgc     2024
Ser Leu Lys Gly Cys Cys Thr Leu Leu Asn Pro Ser Thr Phe Asn Cys
        645                 650                 655

aat ggg atc tct aat aca ttt tga ggtcagccac ttctccagtg gaactctgaa    2078
Asn Gly Ile Ser Asn Thr Phe  *
    660                 665

gtacagatgc tgaattttct gctttggaaa gaaaaaaaaa agaagctatt gtccacagat   2138

tgatgcttcc ataatggaat cagctttaat tgcaaggaat gaagaaaaac aagagtggac   2198

cttcaaagct acaacatttt cctnctnccc ttccttccca ccagcccct               2247


<210> SEQ ID NO 61
<211> LENGTH: 665
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 61

Met Glu Leu Arg Ser Thr Ala Ala Pro Arg Ala Glu Gly Tyr Ser Asn
 1               5                  10                  15

Val Gly Phe Gln Asn Glu Glu Asn Phe Leu Glu Asn Glu Asn Thr Ser
            20                  25                  30

Gly Asn Asn Ser Ile Arg Ser Arg Ala Val Gln Ser Arg Glu His Thr
        35                  40                  45

Asn Thr Lys Gln Asp Glu Glu Gln Val Thr Val Glu Gln Asp Phe Pro
    50                  55                  60

Arg Asn Arg Glu His Met Glu Asp Asp Asp Glu Glu Met Gln Gln Lys
65                  70                  75                  80

Gly Cys Leu Glu Arg Arg Tyr Asp Thr Val Cys Gly Phe Cys Arg Lys
```

-continued

```
                85                  90                  95

His Lys Thr Thr Leu Arg His Ile Ile Trp Gly Ile Leu Leu Ala Gly
            100                 105                 110

Tyr Leu Val Met Val Ile Ser Ala Cys Val Leu Asn Phe His Arg Ala
        115                 120                 125

Leu Pro Leu Phe Val Ile Thr Val Ala Ala Ile Phe Phe Val Val Trp
    130                 135                 140

Asp His Leu Met Ala Lys Tyr Glu His Arg Ile Asp Glu Met Leu Ser
145                 150                 155                 160

Pro Gly Arg Arg Leu Leu Asn Ser His Trp Phe Trp Leu Lys Trp Val
                165                 170                 175

Ile Trp Ser Ser Leu Val Leu Ala Val Ile Phe Trp Leu Ala Phe Asp
            180                 185                 190

Thr Ala Lys Leu Gly Gln Gln Gln Leu Val Ser Phe Gly Gly Leu Ile
        195                 200                 205

Met Tyr Ile Val Leu Leu Phe Leu Phe Ser Lys Tyr Pro Thr Arg Val
    210                 215                 220

Tyr Trp Arg Pro Val Leu Trp Gly Ile Gly Leu Gln Phe Leu Leu Gly
225                 230                 235                 240

Leu Leu Ile Leu Arg Thr Asp Pro Gly Phe Ile Ala Phe Asp Trp Leu
                245                 250                 255

Gly Arg Gln Val Gln Val Leu Pro Ile Val Val Phe Phe Ser Thr Val
            260                 265                 270

Met Ser Met Leu Tyr Tyr Leu Gly Leu Met Gln Trp Ile Ile Arg Lys
        275                 280                 285

Val Gly Trp Ile Met Leu Val Thr Thr Gly Ser Ser Pro Ile Glu Ser
    290                 295                 300

Val Val Ala Ser Gly Asn Ile Phe Val Gly Gln Thr Glu Ser Pro Leu
305                 310                 315                 320

Leu Val Arg Pro Tyr Leu Pro Tyr Ile Thr Lys Ser Glu Leu His Ala
                325                 330                 335

Ile Met Thr Ala Gly Phe Ser Thr Ile Ala Gly Ser Val Leu Gly Ala
            340                 345                 350

Tyr Ile Ser Phe Gly Val Pro Ser Ser His Leu Leu Thr Ala Ser Val
        355                 360                 365

Met Ser Ala Pro Ala Ser Leu Ala Ala Ala Lys Leu Phe Trp Pro Glu
    370                 375                 380

Thr Glu Lys Pro Lys Ile Thr Leu Lys Asn Ala Met Lys Met Glu Ser
385                 390                 395                 400

Gly Asp Ser Gly Asn Leu Leu Glu Ala Ala Thr Gln Gly Ala Ser Ser
                405                 410                 415

Ser Ile Ser Leu Val Ala Asn Ile Ala Val Asn Leu Ile Ala Phe Leu
            420                 425                 430

Ala Leu Leu Ser Phe Met Asn Ser Ala Leu Ser Trp Phe Gly Asn Met
        435                 440                 445

Phe Asp Tyr Pro Gln Leu Ser Phe Glu Leu Ile Cys Ser Tyr Ile Phe
    450                 455                 460

Met Pro Phe Ser Phe Met Met Gly Val Glu Trp Gln Asp Ser Phe Met
465                 470                 475                 480

Val Ala Arg Leu Ile Gly Tyr Lys Thr Phe Phe Asn Glu Phe Val Ala
                485                 490                 495
```

-continued

```
Tyr Glu His Leu Ser Lys Trp Ile His Leu Arg Lys Glu Gly Gly Pro
            500                 505                 510

Lys Phe Val Asn Gly Val Gln Gln Tyr Ile Ser Ile Arg Ser Glu Ile
        515                 520                 525

Ile Ala Thr Tyr Ala Leu Cys Gly Phe Ala Asn Ile Gly Ser Leu Gly
    530                 535                 540

Ile Val Ile Gly Gly Leu Thr Ser Met Ala Pro Ser Arg Lys Arg Asp
545                 550                 555                 560

Ile Ala Ser Gly Ala Val Arg Ala Leu Ile Ala Gly Thr Val Ala Cys
                565                 570                 575

Phe Met Thr Ala Cys Ile Ala Gly Ile Leu Ser Ser Thr Pro Val Asp
            580                 585                 590

Ile Asn Cys His His Val Leu Glu Asn Ala Phe Asn Ser Thr Phe Pro
        595                 600                 605

Gly Asn Thr Thr Lys Val Ile Ala Cys Cys Gln Ser Leu Leu Ser Ser
    610                 615                 620

Thr Val Ala Lys Gly Pro Gly Glu Val Ile Pro Gly Gly Asn His Ser
625                 630                 635                 640

Leu Tyr Ser Leu Lys Gly Cys Cys Thr Leu Leu Asn Pro Ser Thr Phe
                645                 650                 655

Asn Cys Asn Gly Ile Ser Asn Thr Phe
            660                 665


<210> SEQ ID NO 62
<211> LENGTH: 1998
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(1998)

<400> SEQUENCE: 62

atg gag ctg agg agt aca gca gcc ccc aga gct gag ggc tac agc aac       48
Met Glu Leu Arg Ser Thr Ala Ala Pro Arg Ala Glu Gly Tyr Ser Asn
 1               5                  10                  15

gtg ggc ttc cag aat gaa gaa aac ttt ctt gag aac gag aac aca tca       96
Val Gly Phe Gln Asn Glu Glu Asn Phe Leu Glu Asn Glu Asn Thr Ser
            20                  25                  30

gga aac aac tca ata aga agc aga gct gtg caa agc agg gag cac aca      144
Gly Asn Asn Ser Ile Arg Ser Arg Ala Val Gln Ser Arg Glu His Thr
        35                  40                  45

aac acc aaa cag gat gaa gaa cag gtc aca gtt gag cag gat ttt cca      192
Asn Thr Lys Gln Asp Glu Glu Gln Val Thr Val Glu Gln Asp Phe Pro
    50                  55                  60

aga aac aga gaa cac atg gag gat gat gat gag gag atg caa caa aaa      240
Arg Asn Arg Glu His Met Glu Asp Asp Asp Glu Glu Met Gln Gln Lys
65                  70                  75                  80

ggg tgt ttg gaa agg agg tat gac aca gta tgt ggt ttc tgt agg aaa      288
Gly Cys Leu Glu Arg Arg Tyr Asp Thr Val Cys Gly Phe Cys Arg Lys
                85                  90                  95

cac aaa aca act ctt cgg cac atc atc tgg ggc att tta tta gca ggt      336
His Lys Thr Thr Leu Arg His Ile Ile Trp Gly Ile Leu Leu Ala Gly
            100                 105                 110

tat ctg gtt atg gtg att tcg gcc tgt gtg ctg aac ttt cac aga gcc      384
Tyr Leu Val Met Val Ile Ser Ala Cys Val Leu Asn Phe His Arg Ala
        115                 120                 125
```

-continued

```
ctt cct ctt ttt gtg atc acc gtg gct gcc atc ttc ttt gtt gtc tgg      432
Leu Pro Leu Phe Val Ile Thr Val Ala Ala Ile Phe Phe Val Val Trp
    130                 135                 140

gat cac ctg atg gcc aaa tac gaa cat cga att gat gag atg ctg tct      480
Asp His Leu Met Ala Lys Tyr Glu His Arg Ile Asp Glu Met Leu Ser
145                 150                 155                 160

cct ggc aga agg ctt cta aac agc cat tgg ttc tgg ctg aag tgg gtg      528
Pro Gly Arg Arg Leu Leu Asn Ser His Trp Phe Trp Leu Lys Trp Val
                165                 170                 175

atc tgg agc tcc ctg gtc cta gca gtt att ttc tgg ttg gcc ttt gac      576
Ile Trp Ser Ser Leu Val Leu Ala Val Ile Phe Trp Leu Ala Phe Asp
            180                 185                 190

act gcc aaa ttg ggt caa cag cag ctg gtg tcc ttc ggt ggg ctc ata      624
Thr Ala Lys Leu Gly Gln Gln Gln Leu Val Ser Phe Gly Gly Leu Ile
        195                 200                 205

atg tac att gtc ctg tta ttt cta ttc tcc aag tac cca acc aga gtt      672
Met Tyr Ile Val Leu Leu Phe Leu Phe Ser Lys Tyr Pro Thr Arg Val
    210                 215                 220

tac tgg aga cct gtc tta tgg gga atc ggg cta cag ttt ctt ctt ggg      720
Tyr Trp Arg Pro Val Leu Trp Gly Ile Gly Leu Gln Phe Leu Leu Gly
225                 230                 235                 240

ctc ttg att cta agg act gac cct gga ttt ata gct ttt gat tgg ttg      768
Leu Leu Ile Leu Arg Thr Asp Pro Gly Phe Ile Ala Phe Asp Trp Leu
                245                 250                 255

ggc aga caa gtt cag gtc ctg ccg atc gtg gtt ttc ttc agc act gtg      816
Gly Arg Gln Val Gln Val Leu Pro Ile Val Val Phe Phe Ser Thr Val
            260                 265                 270

atg tcc atg ctg tac tac ctg gga ctg atg cag tgg att att aga aag      864
Met Ser Met Leu Tyr Tyr Leu Gly Leu Met Gln Trp Ile Ile Arg Lys
        275                 280                 285

gtt gga tgg atc atg cta gtt act acg gga tca tct cct att gaa tct      912
Val Gly Trp Ile Met Leu Val Thr Thr Gly Ser Ser Pro Ile Glu Ser
    290                 295                 300

gta gtt gct tct ggc aat ata ttt gtt gga caa acg gag tct cct ctg      960
Val Val Ala Ser Gly Asn Ile Phe Val Gly Gln Thr Glu Ser Pro Leu
305                 310                 315                 320

ctg gtc cga cca tat tta cct tac atc acc aag tct gaa ctc cac gcc     1008
Leu Val Arg Pro Tyr Leu Pro Tyr Ile Thr Lys Ser Glu Leu His Ala
                325                 330                 335

atc atg acc gcc ggg ttc tct acc att gct gga agc gtg cta ggt gca     1056
Ile Met Thr Ala Gly Phe Ser Thr Ile Ala Gly Ser Val Leu Gly Ala
            340                 345                 350

tac att tct ttt ggg gtt cca tcc tcc cac ttg tta aca gcg tca gtt     1104
Tyr Ile Ser Phe Gly Val Pro Ser Ser His Leu Leu Thr Ala Ser Val
        355                 360                 365

atg tca gca cct gcg tca ttg gct gct gct aaa ctc ttt tgg cct gag     1152
Met Ser Ala Pro Ala Ser Leu Ala Ala Ala Lys Leu Phe Trp Pro Glu
    370                 375                 380

aca gaa aaa cct aaa ata acc ctc aag aat gcc atg aaa atg gaa agt     1200
Thr Glu Lys Pro Lys Ile Thr Leu Lys Asn Ala Met Lys Met Glu Ser
385                 390                 395                 400

ggt gat tca ggg aat ctt cta gaa gct gca aca cag gga gca tcc tcc     1248
Gly Asp Ser Gly Asn Leu Leu Glu Ala Ala Thr Gln Gly Ala Ser Ser
                405                 410                 415

tcc atc tcc ctg gtg gcc aac atc gct gtg aat ctg att gcc ttc ctg     1296
Ser Ile Ser Leu Val Ala Asn Ile Ala Val Asn Leu Ile Ala Phe Leu
            420                 425                 430
```

-continued

```
gcc ctg ctg tct ttt atg aat tca gcc ctg tcc tgg ttt gga aac atg     1344
Ala Leu Leu Ser Phe Met Asn Ser Ala Leu Ser Trp Phe Gly Asn Met
        435                 440                 445

ttt gac tac cca cag ctg agt ttt gag cta atc tgc tcc tac atc ttc     1392
Phe Asp Tyr Pro Gln Leu Ser Phe Glu Leu Ile Cys Ser Tyr Ile Phe
    450                 455                 460

atg ccc ttt tcc ttc atg atg gga gtg gaa tgg cag gac agc ttt atg     1440
Met Pro Phe Ser Phe Met Met Gly Val Glu Trp Gln Asp Ser Phe Met
465                 470                 475                 480

gtt gcc aga ctc ata ggt tat aag acc ttc ttc aat gaa ttt gtg gct     1488
Val Ala Arg Leu Ile Gly Tyr Lys Thr Phe Phe Asn Glu Phe Val Ala
                485                 490                 495

tat gag cac ctc tca aaa tgg atc cac ttg agg aaa gaa ggt gga ccc     1536
Tyr Glu His Leu Ser Lys Trp Ile His Leu Arg Lys Glu Gly Gly Pro
            500                 505                 510

aaa ttt gta aac ggt gtg cag caa tat ata tca att cgt tct gag ata     1584
Lys Phe Val Asn Gly Val Gln Gln Tyr Ile Ser Ile Arg Ser Glu Ile
        515                 520                 525

atc gcc act tac gct ctc tgt ggt ttt gcc aat atc ggg tcc cta gga     1632
Ile Ala Thr Tyr Ala Leu Cys Gly Phe Ala Asn Ile Gly Ser Leu Gly
    530                 535                 540

atc gtg atc ggc gga ctc aca tcc atg gct cct tcc aga aag cgt gat     1680
Ile Val Ile Gly Gly Leu Thr Ser Met Ala Pro Ser Arg Lys Arg Asp
545                 550                 555                 560

atc gcc tcg ggg gca gtg aga gct ctg att gcg ggg acc gtg gcc tgc     1728
Ile Ala Ser Gly Ala Val Arg Ala Leu Ile Ala Gly Thr Val Ala Cys
                565                 570                 575

ttc atg aca gcc tgc atc gca ggc ata ctc tcc agc act cct gtg gac     1776
Phe Met Thr Ala Cys Ile Ala Gly Ile Leu Ser Ser Thr Pro Val Asp
            580                 585                 590

atc aac tgc cat cac gtt tta gag aat gcc ttc aac tcc act ttc cct     1824
Ile Asn Cys His His Val Leu Glu Asn Ala Phe Asn Ser Thr Phe Pro
        595                 600                 605

gga aac aca acc aag gtg ata gct tgt tgc caa agt ctg ttg agc agc     1872
Gly Asn Thr Thr Lys Val Ile Ala Cys Cys Gln Ser Leu Leu Ser Ser
    610                 615                 620

act gtt gcc aag ggt cct ggt gaa gtc atc cca gga gga aac cac agt     1920
Thr Val Ala Lys Gly Pro Gly Glu Val Ile Pro Gly Gly Asn His Ser
625                 630                 635                 640

ctg tat tct ttg aag ggc tgc tgc aca ttg ttg aat cca tcg acc ttt     1968
Leu Tyr Ser Leu Lys Gly Cys Cys Thr Leu Leu Asn Pro Ser Thr Phe
                645                 650                 655

aac tgc aat ggg atc tct aat aca ttt tga                             1998
Asn Cys Asn Gly Ile Ser Asn Thr Phe  *
            660                 665


<210> SEQ ID NO 63
<211> LENGTH: 427
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 63

Met Glu Val Leu Ile Ser Val Leu Gly Leu Val Val Phe Ile Ala Ile
 1               5                  10                  15

Ala Phe Leu Phe Ser Ser Asn Lys Lys Lys Ile Ser Trp Arg Thr Val
            20                  25                  30
```

-continued

```
Ile Ser Ala Leu Val Leu Gln Val Val Leu Gly Leu Ile Val Leu Lys
        35                  40                  45

Thr Pro Val Gly Arg Trp Val Phe Gly Lys Leu Ala Glu Gly Val Gln
    50                  55                  60

Lys Leu Leu Ala Tyr Ala Asn Glu Gly Ile Asn Phe Val Phe Gly Ser
65                  70                  75                  80

Leu Ala Gly Pro Asp Lys Ile Asp Asp Leu Gly Phe Val Phe Ala Phe
                85                  90                  95

Lys Val Leu Pro Ile Ile Ile Phe Phe Ser Ala Leu Ile Ser Ile Leu
            100                 105                 110

Tyr Tyr Leu Gly Ile Met Gln Val Val Ile Arg Lys Ile Gly Trp Ala
        115                 120                 125

Leu Gln Lys Ala Leu Gly Thr Ser Lys Leu Glu Ser Leu Ser Ala Ala
    130                 135                 140

Ala Asn Ile Phe Leu Gly Gln Thr Glu Ala Pro Leu Val Ile Lys Pro
145                 150                 155                 160

Tyr Leu Gly Lys Leu Thr Arg Ser Glu Leu Phe Thr Val Met Thr Ser
                165                 170                 175

Gly Met Ala Ser Ile Ala Gly Ser Val Leu Gly Ala Tyr Ala Ala Met
            180                 185                 190

Gly Val Pro Pro Glu Tyr Leu Leu Ala Ala Ser Val Leu Ala Ala Pro
        195                 200                 205

Gly Gly Leu Ile Phe Ala Lys Leu Ile Tyr Pro Glu Thr Glu Glu Ser
    210                 215                 220

Gln Glu Lys Glu Glu Glu Ile Leu Lys Leu Glu Glu Glu Glu Lys Lys
225                 230                 235                 240

Asn Phe Phe Glu Ala Leu Ala Asn Gly Ala Leu Ala Gly Leu Lys Val
                245                 250                 255

Ala Leu Asn Val Ala Ala Met Leu Ile Ala Phe Val Ala Leu Ile Ala
            260                 265                 270

Leu Ile Asn Gly Ile Leu Gly Gly Val Gly Gly Leu Phe Gly Tyr Glu
        275                 280                 285

Gly Leu Ser Phe Gln Ser Ile Leu Gly Tyr Leu Phe Ser Pro Leu Ala
    290                 295                 300

Phe Leu Met Gly Val Pro Asp Asn Trp Ser Asp Ala Leu Leu Val Gly
305                 310                 315                 320

Ser Leu Met Gly Thr Lys Leu Ala Val Asn Glu Phe Val Ala Tyr Leu
                325                 330                 335

Asp Leu Ser Lys Tyr Leu Lys Thr Arg Leu Ala Gly Ala Glu Glu Trp
            340                 345                 350

Val Glu Gly Glu Lys Gln Lys Leu Ser Pro Lys Thr Val Ala Ile Ile
        355                 360                 365

Thr Phe Ala Leu Cys Gly Phe Ala Asn Phe Ser Ser Ile Gly Ile Ile
    370                 375                 380

Leu Gly Gly Leu Gly Gly Leu Ala Pro Gly Ser Arg Lys Ser Val Ile
385                 390                 395                 400

Ser Arg Leu Gly Leu Lys Ala Leu Leu Ala Gly Thr Leu Val Ser Leu
                405                 410                 415

Leu Ser Ala Thr Ile Ala Gly Ile Leu Phe Ile
            420                 425
```

-continued

<210> SEQ ID NO 64
<211> LENGTH: 322
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 64

Val Leu Pro Val Ile Ile Phe Phe Ser Ala Leu Ile Ser Ile Leu Tyr
1 5 10 15

Tyr Leu Gly Ile Met Gln Trp Val Ile Arg Lys Ile Gly Trp Leu Leu
20 25 30

Gln Lys Val Leu Gly Thr Ser Lys Val Glu Ser Met Ser Ala Ala Ala
35 40 45

Asn Ile Phe Leu Gly Gln Thr Glu Ala Pro Leu Val Ile Arg Pro Tyr
50 55 60

Leu Glu Lys Met Thr Gln Ser Glu Leu Phe Ala Ile Met Thr Ser Gly
65 70 75 80

Met Ala Thr Val Ala Gly Ser Val Leu Gly Ala Tyr Val Ser Met Gly
85 90 95

Val Pro Ala Glu Tyr Leu Ile Ala Ala Ser Val Met Ser Ala Pro Gly
100 105 110

Gly Leu Ile Phe Ala Lys Leu Ile Tyr Pro Glu Thr Glu Glu Ser Lys
115 120 125

Glu Glu Ser Asp Asp Ser Val Lys Leu Glu Glu Glu Glu Glu Gln Asn
130 135 140

Phe Ile Asp Ala Ala Ala Asn Gly Ala Met Ala Gly Leu Lys Ile Val
145 150 155 160

Leu Asn Ile Ala Ala Met Leu Ile Ala Phe Val Ala Leu Ile Ala Phe
165 170 175

Ile Asn Gly Ile Leu Ser Trp Ile Gly Gly Leu Phe Gly Tyr Glu Gly
180 185 190

Leu Thr Phe Gln Met Ile Phe Gly Tyr Ile Phe Ala Pro Leu Ala Phe
195 200 205

Leu Met Gly Val Pro Trp Asn Asp Glu Ala Met Leu Val Gly Gln Leu
210 215 220

Ile Gly Gln Lys Leu Val Ile Asn Glu Phe Val Ala Tyr Met Asn Leu
225 230 235 240

Ser Gln Tyr Leu Gln Asn Asn Lys Ala Gly Val Trp Val Gly Lys Gln
245 250 255

Trp Leu Ser Pro Arg Thr Glu Ala Ile Val Thr Phe Ala Leu Cys Gly
260 265 270

Phe Ala Asn Phe Ser Ser Ile Gly Ile Leu Ile Gly Gly Leu Gly Gly
275 280 285

Leu Ala Pro Glu Arg Arg Ser Asp Val Ile Ala Arg Leu Gly Leu Arg
290 295 300

Ala Leu Ile Ala Gly Thr Leu Ala Asn Phe Met Ser Ala Cys Ile Ala
305 310 315 320

Gly Ile

<210> SEQ ID NO 65
<211> LENGTH: 61
<212> TYPE: PRT

-continued

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 65

Leu Ile Ser Ile Ile Gly Met Val Val Phe Ile Leu Ile Ala Phe Leu
1 5 10 15

Phe Ser Ser Asn Arg Lys Lys Ile Asn Trp Arg Thr Val Ile Thr Ala
20 25 30

Leu Val Leu Gln Phe Leu Leu Gly Trp Ile Ile Leu Arg Thr Pro Val
35 40 45

Gly Arg Trp Ala Phe Gln Trp Leu Gly Asp Gly Val Gln
50 55 60

<210> SEQ ID NO 66
<211> LENGTH: 77
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 66

Ser Asp Thr Pro Val Asp Ile Asn Cys His His Val Leu Glu Asn Gly
1 5 10 15

Phe Asn Ser Arg Phe Pro Gly Asn Thr Thr Glu Val Ile Ala Cys Cys
20 25 30

Gln Asn Leu Phe Asn Ser Thr Val Ala Lys Gly Pro Asn Asp Val Ile
35 40 45

Pro Gly Gly Asn His Ser Leu Tyr Ala Leu Lys Gly Cys Cys Asn Leu
50 55 60

Leu Asn Pro Pro Thr Phe Asn Cys Asn Trp Ile Pro Asn
65 70 75

<210> SEQ ID NO 67
<211> LENGTH: 103
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 67

Phe Cys Arg Lys His Lys Gln Leu Phe Arg Trp Ile Ile Trp Gly Leu
1 5 10 15

Leu Cys Thr Gly Tyr Leu Ala Phe Ile Leu Ala Ala Cys Ile Leu Asn
20 25 30

Phe Gln Arg Ala Leu Ala Leu Phe Val Ile Thr Cys Val Val Ile Phe
35 40 45

Phe Leu Val Trp Asp Phe Leu Lys Lys Lys Tyr Gly Lys Lys Ile Asp
50 55 60

Arg Cys Leu Lys Pro Gly Arg His Thr Leu Leu Asn His Trp Phe Trp
65 70 75 80

Leu Lys Trp Val Ile Trp Gly Ala Ala Ile Leu Gly Leu Ile Leu Trp
85 90 95

Leu Ala Leu Asp Thr Ala Gln
100

<210> SEQ ID NO 68

-continued

<211> LENGTH: 91
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 68

Met Glu Leu Arg Pro Thr Ala Ala Pro Arg Asp Glu Gly Tyr Ser Asn
1 5 10 15

Glu Gly Phe Gln Asn Glu Asp Asn Phe Leu Glu Asn Gln Asn Thr Pro
20 25 30

Gly Asn Asn Thr Ile Arg Asn Arg Ala Val Gln Ser Gly Glu His Gly
35 40 45

His Thr Lys Gln Asp Asp Arg Gln Ile Thr Ile Glu Gln Asp Pro Pro
50 55 60

Gly Asn Arg Glu His Pro Glu Asp Asp Asp Glu Asp Glu His Gln Lys
65 70 75 80

Gly Cys Leu Glu Arg Arg Tyr Asp Thr Ile Cys
85 90

<210> SEQ ID NO 69
<211> LENGTH: 2473
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (88)...(2241)

<400> SEQUENCE: 69

gaaatgatgg agtaagagac tcttttctaa gcaactcaag tttgcagtga ttcaggccta 60

cttctgaaga gacagccttt tatctca atg aat gac aca gaa aaa cca gca gat 114
Met Asn Asp Thr Glu Lys Pro Ala Asp
1 5

act ccc tct gag gaa gag gac ttt ggt gat cca agg aca tat gac cca 162
Thr Pro Ser Glu Glu Glu Asp Phe Gly Asp Pro Arg Thr Tyr Asp Pro
10 15 20 25

gat ttc aag ggg cct gtt gcc aac agg agt tgt aca gat gtt ctg tgc 210
Asp Phe Lys Gly Pro Val Ala Asn Arg Ser Cys Thr Asp Val Leu Cys
30 35 40

tgt atg atc ttc cta ctg tgt att att ggc tac att gtt tta gga ctt 258
Cys Met Ile Phe Leu Leu Cys Ile Ile Gly Tyr Ile Val Leu Gly Leu
45 50 55

gtg gcc tgg gta cat ggg gac ccc aga aga gca gcc tat cct aca gac 306
Val Ala Trp Val His Gly Asp Pro Arg Arg Ala Ala Tyr Pro Thr Asp
60 65 70

agc cag ggc cac ttt tgt ggc cag aag ggc act ccc aat gag aac aag 354
Ser Gln Gly His Phe Cys Gly Gln Lys Gly Thr Pro Asn Glu Asn Lys
75 80 85

acc att ttg ttt tac ttt aac ctg tta cgc tgt acc agt ccc tcc gtg 402
Thr Ile Leu Phe Tyr Phe Asn Leu Leu Arg Cys Thr Ser Pro Ser Val
90 95 100 105

ttg cta aac cta cag tgc cct acc aca cag atc tgt gtc tcc aag tgc 450
Leu Leu Asn Leu Gln Cys Pro Thr Thr Gln Ile Cys Val Ser Lys Cys
110 115 120

cca gaa aaa ttt tta acc tat gtg gaa atg caa ctt ttg tac aca aaa 498
Pro Glu Lys Phe Leu Thr Tyr Val Glu Met Gln Leu Leu Tyr Thr Lys
125 130 135

gac aaa agc tac tgg gaa gac tac cgt cag ttc tgt aag acc act gct 546

-continued

```
Asp Lys Ser Tyr Trp Glu Asp Tyr Arg Gln Phe Cys Lys Thr Thr Ala
        140                 145                 150

aag cct gtg aag tct ctc aca cag ctt tta ctg gat gat gat tgt cca      594
Lys Pro Val Lys Ser Leu Thr Gln Leu Leu Leu Asp Asp Asp Cys Pro
    155                 160                 165

aca gcg att ttt ccc agc aaa cct ttt ctc cag aga tgt ttc cct gac      642
Thr Ala Ile Phe Pro Ser Lys Pro Phe Leu Gln Arg Cys Phe Pro Asp
170                 175                 180                 185

ttc tct acc aaa aat ggc act tta aca ata gga agt aag atg atg ttt      690
Phe Ser Thr Lys Asn Gly Thr Leu Thr Ile Gly Ser Lys Met Met Phe
                190                 195                 200

caa gat gga aat gga ggg aca aga agt gtt gta gaa ctc ggg att gct      738
Gln Asp Gly Asn Gly Gly Thr Arg Ser Val Val Glu Leu Gly Ile Ala
            205                 210                 215

gca aat ggt atc aat aaa ctt ctt gat gca aag tca ctt gga ttg aaa      786
Ala Asn Gly Ile Asn Lys Leu Leu Asp Ala Lys Ser Leu Gly Leu Lys
        220                 225                 230

gtg ttt gaa gac tat gca aga act tgg tat tgg att ctt att ggc ctg      834
Val Phe Glu Asp Tyr Ala Arg Thr Trp Tyr Trp Ile Leu Ile Gly Leu
    235                 240                 245

acg atc gcc atg gtc ctt agt tgg ata ttt ttg ata ctt ctg agg ttc      882
Thr Ile Ala Met Val Leu Ser Trp Ile Phe Leu Ile Leu Leu Arg Phe
250                 255                 260                 265

ata gct gga tgc ctc ttc tgg gtc ttc atg att ggt gtg att gga att      930
Ile Ala Gly Cys Leu Phe Trp Val Phe Met Ile Gly Val Ile Gly Ile
                270                 275                 280

ata ggt tat gga ata tgg cac tgt tac cag cag tac acc aat ctt cag      978
Ile Gly Tyr Gly Ile Trp His Cys Tyr Gln Gln Tyr Thr Asn Leu Gln
            285                 290                 295

gaa cgc cca agt tct gta tta act atc tat gac atc ggg att cag act     1026
Glu Arg Pro Ser Ser Val Leu Thr Ile Tyr Asp Ile Gly Ile Gln Thr
        300                 305                 310

aac ata agc atg tac ttt gaa ctg caa caa aca tgg ttc aca ttt atg     1074
Asn Ile Ser Met Tyr Phe Glu Leu Gln Gln Thr Trp Phe Thr Phe Met
    315                 320                 325

ata ata ctc tgc atc att gaa gtg att gtc atc ctc atg ctg atc ttc     1122
Ile Ile Leu Cys Ile Ile Glu Val Ile Val Ile Leu Met Leu Ile Phe
330                 335                 340                 345

ctc agg aat cga atc cga gtc gcc att atc ctg ctg aag gaa gga agc     1170
Leu Arg Asn Arg Ile Arg Val Ala Ile Ile Leu Leu Lys Glu Gly Ser
                350                 355                 360

aaa gcc att gga tat gtt cct agt aca tta gtc tat cca gct tta act     1218
Lys Ala Ile Gly Tyr Val Pro Ser Thr Leu Val Tyr Pro Ala Leu Thr
            365                 370                 375

ttc att ttg ctc tca atc tgc att tgc tac tgg gtc gtg aca gca gtt     1266
Phe Ile Leu Leu Ser Ile Cys Ile Cys Tyr Trp Val Val Thr Ala Val
        380                 385                 390

ttc ttg gcg aca tcg ggg gta cct gta tac aaa gtc ata gct cca ggg     1314
Phe Leu Ala Thr Ser Gly Val Pro Val Tyr Lys Val Ile Ala Pro Gly
    395                 400                 405

ggg cat tgt ata cat gaa aat caa acc tgt gac cca gag att ttt aat     1362
Gly His Cys Ile His Glu Asn Gln Thr Cys Asp Pro Glu Ile Phe Asn
410                 415                 420                 425

aca act gaa att gcc aaa gct tgc cct ggg gct ctg tgt aac ttt gct     1410
Thr Thr Glu Ile Ala Lys Ala Cys Pro Gly Ala Leu Cys Asn Phe Ala
                430                 435                 440

ttc tat ggt gga aag agc ttg tac cat cag tac atc cct acc ttc cat     1458
```

-continued

```
Phe Tyr Gly Gly Lys Ser Leu Tyr His Gln Tyr Ile Pro Thr Phe His
            445                 450                 455

gta tac aac tta ttt gtc ttt ctc tgg ctt ata aac ttc gtc att gca     1506
Val Tyr Asn Leu Phe Val Phe Leu Trp Leu Ile Asn Phe Val Ile Ala
        460                 465                 470

tta ggt cag tgc gcc ctt gct ggt gca ttc gct act tat tac tgg gcc     1554
Leu Gly Gln Cys Ala Leu Ala Gly Ala Phe Ala Thr Tyr Tyr Trp Ala
    475                 480                 485

atg aaa aaa cct gat gac atc cca cga tat cca ctt ttt act gca ttt     1602
Met Lys Lys Pro Asp Asp Ile Pro Arg Tyr Pro Leu Phe Thr Ala Phe
490                 495                 500                 505

gga cga gcc ata cga tat cac aca gga tcc cta gca ttt gga tct tta     1650
Gly Arg Ala Ile Arg Tyr His Thr Gly Ser Leu Ala Phe Gly Ser Leu
                510                 515                 520

att att gca tta att caa atg ttt aaa att gta cta gaa tac ttg gac     1698
Ile Ile Ala Leu Ile Gln Met Phe Lys Ile Val Leu Glu Tyr Leu Asp
            525                 530                 535

cac cgt ctt aaa cgt acc cag aac aca ttg tct aaa ttc cta cag tgc     1746
His Arg Leu Lys Arg Thr Gln Asn Thr Leu Ser Lys Phe Leu Gln Cys
        540                 545                 550

tgc ctg aga tgc tgc ttc tgg tgt ttg gaa aat gca ata aag ttt tta     1794
Cys Leu Arg Cys Cys Phe Trp Cys Leu Glu Asn Ala Ile Lys Phe Leu
    555                 560                 565

aac aga aat gcc tat att atg att gca ata tat ggc aga aac ttc tgc     1842
Asn Arg Asn Ala Tyr Ile Met Ile Ala Ile Tyr Gly Arg Asn Phe Cys
570                 575                 580                 585

agg tca gca aaa gat gct ttc aat ctg ctg atg aga aat gtt ttg aaa     1890
Arg Ser Ala Lys Asp Ala Phe Asn Leu Leu Met Arg Asn Val Leu Lys
                590                 595                 600

gtt gca gtt aca gat gaa gtt aca tac ttt gta tta ttc ctg ggg aaa     1938
Val Ala Val Thr Asp Glu Val Thr Tyr Phe Val Leu Phe Leu Gly Lys
            605                 610                 615

ctt cta gtt gct gga agt ata ggt gtt ctg gcc ttc cta ttc ttc aca     1986
Leu Leu Val Ala Gly Ser Ile Gly Val Leu Ala Phe Leu Phe Phe Thr
        620                 625                 630

caa aga ctg cca gtg att gca caa gga cca gca tct tta aat tac tac     2034
Gln Arg Leu Pro Val Ile Ala Gln Gly Pro Ala Ser Leu Asn Tyr Tyr
    635                 640                 645

tgg gta cct ttg ctg aca gtc att ttt ggg tct tac ctg att gca cat     2082
Trp Val Pro Leu Leu Thr Val Ile Phe Gly Ser Tyr Leu Ile Ala His
650                 655                 660                 665

ggg ttc ttc agc gtc tat gca atg tgt gtt gaa aca att ttc atc tgc     2130
Gly Phe Phe Ser Val Tyr Ala Met Cys Val Glu Thr Ile Phe Ile Cys
                670                 675                 680

ttc ttg gaa gat tta gaa aga aat gat ggt tct act gca aga cct tat     2178
Phe Leu Glu Asp Leu Glu Arg Asn Asp Gly Ser Thr Ala Arg Pro Tyr
            685                 690                 695

tat gtg agt caa cct ttg ctg aag att ttc cag gag gaa aat cca caa     2226
Tyr Val Ser Gln Pro Leu Leu Lys Ile Phe Gln Glu Glu Asn Pro Gln
        700                 705                 710

act agg aag cag tag aagagcaaac tggtcgtcct acagctgtgt gttacctttt     2281
Thr Arg Lys Gln  *
    715

ctccatctgc tgtgtctgtg caacatttgt ttcataagtg ctttgtgttt agcaacactg   2341

tattcacgac cttgttggct tgcatttgca tgttttatac caaagcttat actgtactat   2401

gtgaagccat cagaagtcgc aagggaattg ttaataacat aaaacatttt tatactaaaa   2461
```

-continued

aaaaaaaaaa aa 2473

<210> SEQ ID NO 70
<211> LENGTH: 717
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 70

Met Asn Asp Thr Glu Lys Pro Ala Asp Thr Pro Ser Glu Glu Glu Asp
1 5 10 15

Phe Gly Asp Pro Arg Thr Tyr Asp Pro Asp Phe Lys Gly Pro Val Ala
20 25 30

Asn Arg Ser Cys Thr Asp Val Leu Cys Cys Met Ile Phe Leu Leu Cys
35 40 45

Ile Ile Gly Tyr Ile Val Leu Gly Leu Val Ala Trp Val His Gly Asp
50 55 60

Pro Arg Arg Ala Ala Tyr Pro Thr Asp Ser Gln Gly His Phe Cys Gly
65 70 75 80

Gln Lys Gly Thr Pro Asn Glu Asn Lys Thr Ile Leu Phe Tyr Phe Asn
85 90 95

Leu Leu Arg Cys Thr Ser Pro Ser Val Leu Leu Asn Leu Gln Cys Pro
100 105 110

Thr Thr Gln Ile Cys Val Ser Lys Cys Pro Glu Lys Phe Leu Thr Tyr
115 120 125

Val Glu Met Gln Leu Leu Tyr Thr Lys Asp Lys Ser Tyr Trp Glu Asp
130 135 140

Tyr Arg Gln Phe Cys Lys Thr Thr Ala Lys Pro Val Lys Ser Leu Thr
145 150 155 160

Gln Leu Leu Leu Asp Asp Asp Cys Pro Thr Ala Ile Phe Pro Ser Lys
165 170 175

Pro Phe Leu Gln Arg Cys Phe Pro Asp Phe Ser Thr Lys Asn Gly Thr
180 185 190

Leu Thr Ile Gly Ser Lys Met Met Phe Gln Asp Gly Asn Gly Gly Thr
195 200 205

Arg Ser Val Val Glu Leu Gly Ile Ala Ala Asn Gly Ile Asn Lys Leu
210 215 220

Leu Asp Ala Lys Ser Leu Gly Leu Lys Val Phe Glu Asp Tyr Ala Arg
225 230 235 240

Thr Trp Tyr Trp Ile Leu Ile Gly Leu Thr Ile Ala Met Val Leu Ser
245 250 255

Trp Ile Phe Leu Ile Leu Leu Arg Phe Ile Ala Gly Cys Leu Phe Trp
260 265 270

Val Phe Met Ile Gly Val Ile Gly Ile Ile Gly Tyr Gly Ile Trp His
275 280 285

Cys Tyr Gln Gln Tyr Thr Asn Leu Gln Glu Arg Pro Ser Ser Val Leu
290 295 300

Thr Ile Tyr Asp Ile Gly Ile Gln Thr Asn Ile Ser Met Tyr Phe Glu
305 310 315 320

Leu Gln Gln Thr Trp Phe Thr Phe Met Ile Ile Leu Cys Ile Ile Glu
325 330 335

Val Ile Val Ile Leu Met Leu Ile Phe Leu Arg Asn Arg Ile Arg Val
340 345 350

-continued

```
Ala Ile Ile Leu Leu Lys Glu Gly Ser Lys Ala Ile Gly Tyr Val Pro
        355                 360                 365

Ser Thr Leu Val Tyr Pro Ala Leu Thr Phe Ile Leu Leu Ser Ile Cys
    370                 375                 380

Ile Cys Tyr Trp Val Val Thr Ala Val Phe Leu Ala Thr Ser Gly Val
385                 390                 395                 400

Pro Val Tyr Lys Val Ile Ala Pro Gly Gly His Cys Ile His Glu Asn
                405                 410                 415

Gln Thr Cys Asp Pro Glu Ile Phe Asn Thr Thr Glu Ile Ala Lys Ala
            420                 425                 430

Cys Pro Gly Ala Leu Cys Asn Phe Ala Phe Tyr Gly Gly Lys Ser Leu
        435                 440                 445

Tyr His Gln Tyr Ile Pro Thr Phe His Val Tyr Asn Leu Phe Val Phe
    450                 455                 460

Leu Trp Leu Ile Asn Phe Val Ile Ala Leu Gly Gln Cys Ala Leu Ala
465                 470                 475                 480

Gly Ala Phe Ala Thr Tyr Tyr Trp Ala Met Lys Lys Pro Asp Asp Ile
                485                 490                 495

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

Thr Gly Ser Leu Ala Phe Gly Ser Leu Ile Ile Ala Leu Ile Gln Met
        515                 520                 525

Phe Lys Ile Val Leu Glu Tyr Leu Asp His Arg Leu Lys Arg Thr Gln
    530                 535                 540

Asn Thr Leu Ser Lys Phe Leu Gln Cys Cys Leu Arg Cys Cys Phe Trp
545                 550                 555                 560

Cys Leu Glu Asn Ala Ile Lys Phe Leu Asn Arg Asn Ala Tyr Ile Met
                565                 570                 575

Ile Ala Ile Tyr Gly Arg Asn Phe Cys Arg Ser Ala Lys Asp Ala Phe
            580                 585                 590

Asn Leu Leu Met Arg Asn Val Leu Lys Val Ala Val Thr Asp Glu Val
        595                 600                 605

Thr Tyr Phe Val Leu Phe Leu Gly Lys Leu Leu Val Ala Gly Ser Ile
    610                 615                 620

Gly Val Leu Ala Phe Leu Phe Phe Thr Gln Arg Leu Pro Val Ile Ala
625                 630                 635                 640

Gln Gly Pro Ala Ser Leu Asn Tyr Tyr Trp Val Pro Leu Leu Thr Val
                645                 650                 655

Ile Phe Gly Ser Tyr Leu Ile Ala His Gly Phe Phe Ser Val Tyr Ala
            660                 665                 670

Met Cys Val Glu Thr Ile Phe Ile Cys Phe Leu Glu Asp Leu Glu Arg
        675                 680                 685

Asn Asp Gly Ser Thr Ala Arg Pro Tyr Tyr Val Ser Gln Pro Leu Leu
    690                 695                 700

Lys Ile Phe Gln Glu Glu Asn Pro Gln Thr Arg Lys Gln
705                 710                 715
```

<210> SEQ ID NO 71
<211> LENGTH: 2154
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:

-continued

```
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(2154)

<400> SEQUENCE: 71

atg aat gac aca gaa aaa cca gca gat act ccc tct gag gaa gag gac       48
Met Asn Asp Thr Glu Lys Pro Ala Asp Thr Pro Ser Glu Glu Glu Asp
 1               5                  10                  15

ttt ggt gat cca agg aca tat gac cca gat ttc aag ggg cct gtt gcc       96
Phe Gly Asp Pro Arg Thr Tyr Asp Pro Asp Phe Lys Gly Pro Val Ala
            20                  25                  30

aac agg agt tgt aca gat gtt ctg tgc tgt atg atc ttc cta ctg tgt      144
Asn Arg Ser Cys Thr Asp Val Leu Cys Cys Met Ile Phe Leu Leu Cys
        35                  40                  45

att att ggc tac att gtt tta gga ctt gtg gcc tgg gta cat ggg gac      192
Ile Ile Gly Tyr Ile Val Leu Gly Leu Val Ala Trp Val His Gly Asp
    50                  55                  60

ccc aga aga gca gcc tat cct aca gac agc cag ggc cac ttt tgt ggc      240
Pro Arg Arg Ala Ala Tyr Pro Thr Asp Ser Gln Gly His Phe Cys Gly
65                  70                  75                  80

cag aag ggc act ccc aat gag aac aag acc att ttg ttt tac ttt aac      288
Gln Lys Gly Thr Pro Asn Glu Asn Lys Thr Ile Leu Phe Tyr Phe Asn
                85                  90                  95

ctg tta cgc tgt acc agt ccc tcc gtg ttg cta aac cta cag tgc cct      336
Leu Leu Arg Cys Thr Ser Pro Ser Val Leu Leu Asn Leu Gln Cys Pro
            100                 105                 110

acc aca cag atc tgt gtc tcc aag tgc cca gaa aaa ttt tta acc tat      384
Thr Thr Gln Ile Cys Val Ser Lys Cys Pro Glu Lys Phe Leu Thr Tyr
        115                 120                 125

gtg gaa atg caa ctt ttg tac aca aaa gac aaa agc tac tgg gaa gac      432
Val Glu Met Gln Leu Leu Tyr Thr Lys Asp Lys Ser Tyr Trp Glu Asp
    130                 135                 140

tac cgt cag ttc tgt aag acc act gct aag cct gtg aag tct ctc aca      480
Tyr Arg Gln Phe Cys Lys Thr Thr Ala Lys Pro Val Lys Ser Leu Thr
145                 150                 155                 160

cag ctt tta ctg gat gat gat tgt cca aca gcg att ttt ccc agc aaa      528
Gln Leu Leu Leu Asp Asp Asp Cys Pro Thr Ala Ile Phe Pro Ser Lys
                165                 170                 175

cct ttt ctc cag aga tgt ttc cct gac ttc tct acc aaa aat ggc act      576
Pro Phe Leu Gln Arg Cys Phe Pro Asp Phe Ser Thr Lys Asn Gly Thr
            180                 185                 190

tta aca ata gga agt aag atg atg ttt caa gat gga aat gga ggg aca      624
Leu Thr Ile Gly Ser Lys Met Met Phe Gln Asp Gly Asn Gly Gly Thr
        195                 200                 205

aga agt gtt gta gaa ctc ggg att gct gca aat ggt atc aat aaa ctt      672
Arg Ser Val Val Glu Leu Gly Ile Ala Ala Asn Gly Ile Asn Lys Leu
    210                 215                 220

ctt gat gca aag tca ctt gga ttg aaa gtg ttt gaa gac tat gca aga      720
Leu Asp Ala Lys Ser Leu Gly Leu Lys Val Phe Glu Asp Tyr Ala Arg
225                 230                 235                 240

act tgg tat tgg att ctt att ggc ctg acg atc gcc atg gtc ctt agt      768
Thr Trp Tyr Trp Ile Leu Ile Gly Leu Thr Ile Ala Met Val Leu Ser
                245                 250                 255

tgg ata ttt ttg ata ctt ctg agg ttc ata gct gga tgc ctc ttc tgg      816
Trp Ile Phe Leu Ile Leu Leu Arg Phe Ile Ala Gly Cys Leu Phe Trp
            260                 265                 270

gtc ttc atg att ggt gtg att gga att ata ggt tat gga ata tgg cac      864
Val Phe Met Ile Gly Val Ile Gly Ile Ile Gly Tyr Gly Ile Trp His
        275                 280                 285
```

-continued

```
tgt tac cag cag tac acc aat ctt cag gaa cgc cca agt tct gta tta      912
Cys Tyr Gln Gln Tyr Thr Asn Leu Gln Glu Arg Pro Ser Ser Val Leu
    290                 295                 300

act atc tat gac atc ggg att cag act aac ata agc atg tac ttt gaa      960
Thr Ile Tyr Asp Ile Gly Ile Gln Thr Asn Ile Ser Met Tyr Phe Glu
305                 310                 315                 320

ctg caa caa aca tgg ttc aca ttt atg ata ata ctc tgc atc att gaa     1008
Leu Gln Gln Thr Trp Phe Thr Phe Met Ile Ile Leu Cys Ile Ile Glu
                325                 330                 335

gtg att gtc atc ctc atg ctg atc ttc ctc agg aat cga atc cga gtc     1056
Val Ile Val Ile Leu Met Leu Ile Phe Leu Arg Asn Arg Ile Arg Val
            340                 345                 350

gcc att atc ctg ctg aag gaa gga agc aaa gcc att gga tat gtt cct     1104
Ala Ile Ile Leu Leu Lys Glu Gly Ser Lys Ala Ile Gly Tyr Val Pro
        355                 360                 365

agt aca tta gtc tat cca gct tta act ttc att ttg ctc tca atc tgc     1152
Ser Thr Leu Val Tyr Pro Ala Leu Thr Phe Ile Leu Leu Ser Ile Cys
    370                 375                 380

att tgc tac tgg gtc gtg aca gca gtt ttc ttg gcg aca tcg ggg gta     1200
Ile Cys Tyr Trp Val Val Thr Ala Val Phe Leu Ala Thr Ser Gly Val
385                 390                 395                 400

cct gta tac aaa gtc ata gct cca ggg ggg cat tgt ata cat gaa aat     1248
Pro Val Tyr Lys Val Ile Ala Pro Gly Gly His Cys Ile His Glu Asn
                405                 410                 415

caa acc tgt gac cca gag att ttt aat aca act gaa att gcc aaa gct     1296
Gln Thr Cys Asp Pro Glu Ile Phe Asn Thr Thr Glu Ile Ala Lys Ala
            420                 425                 430

tgc cct ggg gct ctg tgt aac ttt gct ttc tat ggt gga aag agc ttg     1344
Cys Pro Gly Ala Leu Cys Asn Phe Ala Phe Tyr Gly Gly Lys Ser Leu
        435                 440                 445

tac cat cag tac atc cct acc ttc cat gta tac aac tta ttt gtc ttt     1392
Tyr His Gln Tyr Ile Pro Thr Phe His Val Tyr Asn Leu Phe Val Phe
    450                 455                 460

ctc tgg ctt ata aac ttc gtc att gca tta ggt cag tgc gcc ctt gct     1440
Leu Trp Leu Ile Asn Phe Val Ile Ala Leu Gly Gln Cys Ala Leu Ala
465                 470                 475                 480

ggt gca ttc gct act tat tac tgg gcc atg aaa aaa cct gat gac atc     1488
Gly Ala Phe Ala Thr Tyr Tyr Trp Ala Met Lys Lys Pro Asp Asp Ile
                485                 490                 495

cca cga tat cca ctt ttt act gca ttt gga cga gcc ata cga tat cac     1536
Pro Arg Tyr Pro Leu Phe Thr Ala Phe Gly Arg Ala Ile Arg Tyr His
            500                 505                 510

aca gga tcc cta gca ttt gga tct tta att att gca tta att caa atg     1584
Thr Gly Ser Leu Ala Phe Gly Ser Leu Ile Ile Ala Leu Ile Gln Met
        515                 520                 525

ttt aaa att gta cta gaa tac ttg gac cac cgt ctt aaa cgt acc cag     1632
Phe Lys Ile Val Leu Glu Tyr Leu Asp His Arg Leu Lys Arg Thr Gln
    530                 535                 540

aac aca ttg tct aaa ttc cta cag tgc tgc ctg aga tgc tgc ttc tgg     1680
Asn Thr Leu Ser Lys Phe Leu Gln Cys Cys Leu Arg Cys Cys Phe Trp
545                 550                 555                 560

tgt ttg gaa aat gca ata aag ttt tta aac aga aat gcc tat att atg     1728
Cys Leu Glu Asn Ala Ile Lys Phe Leu Asn Arg Asn Ala Tyr Ile Met
                565                 570                 575

att gca ata tat ggc aga aac ttc tgc agg tca gca aaa gat gct ttc     1776
Ile Ala Ile Tyr Gly Arg Asn Phe Cys Arg Ser Ala Lys Asp Ala Phe
            580                 585                 590
```

-continued

```
aat ctg ctg atg aga aat gtt ttg aaa gtt gca gtt aca gat gaa gtt     1824
Asn Leu Leu Met Arg Asn Val Leu Lys Val Ala Val Thr Asp Glu Val
        595                 600                 605

aca tac ttt gta tta ttc ctg ggg aaa ctt cta gtt gct gga agt ata     1872
Thr Tyr Phe Val Leu Phe Leu Gly Lys Leu Leu Val Ala Gly Ser Ile
    610                 615                 620

ggt gtt ctg gcc ttc cta ttc ttc aca caa aga ctg cca gtg att gca     1920
Gly Val Leu Ala Phe Leu Phe Phe Thr Gln Arg Leu Pro Val Ile Ala
625                 630                 635                 640

caa gga cca gca tct tta aat tac tac tgg gta cct ttg ctg aca gtc     1968
Gln Gly Pro Ala Ser Leu Asn Tyr Tyr Trp Val Pro Leu Leu Thr Val
                645                 650                 655

att ttt ggg tct tac ctg att gca cat ggg ttc ttc agc gtc tat gca     2016
Ile Phe Gly Ser Tyr Leu Ile Ala His Gly Phe Phe Ser Val Tyr Ala
            660                 665                 670

atg tgt gtt gaa aca att ttc atc tgc ttc ttg gaa gat tta gaa aga     2064
Met Cys Val Glu Thr Ile Phe Ile Cys Phe Leu Glu Asp Leu Glu Arg
        675                 680                 685

aat gat ggt tct act gca aga cct tat tat gtg agt caa cct ttg ctg     2112
Asn Asp Gly Ser Thr Ala Arg Pro Tyr Tyr Val Ser Gln Pro Leu Leu
    690                 695                 700

aag att ttc cag gag gaa aat cca caa act agg aag cag tag             2154
Lys Ile Phe Gln Glu Glu Asn Pro Gln Thr Arg Lys Gln  *
705                 710                 715


<210> SEQ ID NO 72
<211> LENGTH: 2233
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (110)...(2071)

<400> SEQUENCE: 72

cccacgcgtc cgccagcccc ggccccggcc ccggctcgcg ggcgctgcgt ctccgcgtac     60

aggaggcggc ggcggctccc agtcaccggc ccccgccggc gagcgcacg atg cac tgc    118
Met His Cys
 1

ctg ggc gcc gag tac ctg gtt tct gca gaa gga gcc cct agg caa agg      166
Leu Gly Ala Glu Tyr Leu Val Ser Ala Glu Gly Ala Pro Arg Gln Arg
     5                  10                  15

gag tgg cga ccc cag att tat agg aaa tgc aca gat acg gca tgg tta      214
Glu Trp Arg Pro Gln Ile Tyr Arg Lys Cys Thr Asp Thr Ala Trp Leu
 20                  25                  30                  35

ttc ctg ttc ttt ctc ttt tgg act ggt ttg gtg ttt atc atg ggc tac      262
Phe Leu Phe Phe Leu Phe Trp Thr Gly Leu Val Phe Ile Met Gly Tyr
                 40                  45                  50

tcg gtg gtg gct gga gcc gcg gga aga ctc ctc ttt ggc tat gac agc      310
Ser Val Val Ala Gly Ala Ala Gly Arg Leu Leu Phe Gly Tyr Asp Ser
             55                  60                  65

ttt ggc aac atg tgt ggc aag aag aac tcc ccc gtg gaa ggg gcc cct      358
Phe Gly Asn Met Cys Gly Lys Lys Asn Ser Pro Val Glu Gly Ala Pro
         70                  75                  80

ctt tca ggg cag gac atg acc cta aaa aaa cac gtg ttc ttt atg aat      406
Leu Ser Gly Gln Asp Met Thr Leu Lys Lys His Val Phe Phe Met Asn
     85                  90                  95

tcc tgc aac ctg gaa gtc aaa ggt acg cag ctc aac cgc atg gcc ctc      454
Ser Cys Asn Leu Glu Val Lys Gly Thr Gln Leu Asn Arg Met Ala Leu
```

-continued

```
100                 105                 110                 115

tgt gta tcc aac tgc cct gaa gag cag ctt gac tcc ctg gaa gag gtc      502
Cys Val Ser Asn Cys Pro Glu Glu Gln Leu Asp Ser Leu Glu Glu Val
                120                 125                 130

cag ttc ttt gca aac acc agt ggg tcc ttc ctg tgt gtt tat agt ttg      550
Gln Phe Phe Ala Asn Thr Ser Gly Ser Phe Leu Cys Val Tyr Ser Leu
            135                 140                 145

aat tcc ttc aac tat acc cac agt cca aaa gca gac tca ctg tgt ccc      598
Asn Ser Phe Asn Tyr Thr His Ser Pro Lys Ala Asp Ser Leu Cys Pro
        150                 155                 160

agg cta cca gtt cct cca agc aag tca ttt ccc tta ttt aac cga tgt      646
Arg Leu Pro Val Pro Pro Ser Lys Ser Phe Pro Leu Phe Asn Arg Cys
    165                 170                 175

gtc cct caa aca cct gag tgc tac tcc cta ttt gca tct gtt ttg ata      694
Val Pro Gln Thr Pro Glu Cys Tyr Ser Leu Phe Ala Ser Val Leu Ile
180                 185                 190                 195

aat gat gtt gac acc ctc cac cga att cta agt gga atc atg tcg gga      742
Asn Asp Val Asp Thr Leu His Arg Ile Leu Ser Gly Ile Met Ser Gly
                200                 205                 210

aga gat aca atc ctt ggc ctg tgt atc ctc gca tta gcc ttg tct ttg      790
Arg Asp Thr Ile Leu Gly Leu Cys Ile Leu Ala Leu Ala Leu Ser Leu
            215                 220                 225

gcc atg atg ttt acc ttc aga ttc atc acc acc ctt ctg gtt cac att      838
Ala Met Met Phe Thr Phe Arg Phe Ile Thr Thr Leu Leu Val His Ile
        230                 235                 240

ttc att tca ttg gtt att ttg gga ttg ttg ttt gtc tgc ggt gtt tta      886
Phe Ile Ser Leu Val Ile Leu Gly Leu Leu Phe Val Cys Gly Val Leu
    245                 250                 255

tgg tgg ctg tat tat gac tat acc aac gac ctc agc ata gaa ttg gac      934
Trp Trp Leu Tyr Tyr Asp Tyr Thr Asn Asp Leu Ser Ile Glu Leu Asp
260                 265                 270                 275

aca gaa agg gaa aat atg aag tgc gtg ctg ggg ttt gct atc gta tcc      982
Thr Glu Arg Glu Asn Met Lys Cys Val Leu Gly Phe Ala Ile Val Ser
                280                 285                 290

aca ggc atc acg gca gtg ctg ctc gtc ttg att ttt gtt ctc aga aag     1030
Thr Gly Ile Thr Ala Val Leu Leu Val Leu Ile Phe Val Leu Arg Lys
            295                 300                 305

aga ata aaa ttg aca gtt gag ctt ttc caa atc aca aat aaa gcc atc     1078
Arg Ile Lys Leu Thr Val Glu Leu Phe Gln Ile Thr Asn Lys Ala Ile
        310                 315                 320

agc agt gct ccc ttc ctg ctg ttc cag cca ctg tgg aca ttt gcc atc     1126
Ser Ser Ala Pro Phe Leu Leu Phe Gln Pro Leu Trp Thr Phe Ala Ile
    325                 330                 335

ctc att ttc ttc tgg gtc ctc tgg gtg gct gtg ctg ctg agc ctg gga     1174
Leu Ile Phe Phe Trp Val Leu Trp Val Ala Val Leu Leu Ser Leu Gly
340                 345                 350                 355

act gca gga gct gcc cag gtt atg gaa ggc ggc caa gtg gaa tat aag     1222
Thr Ala Gly Ala Ala Gln Val Met Glu Gly Gly Gln Val Glu Tyr Lys
                360                 365                 370

ccc ctt tcg ggc att cgg tac atg tgg tcg tac cat tta att ggc ctc     1270
Pro Leu Ser Gly Ile Arg Tyr Met Trp Ser Tyr His Leu Ile Gly Leu
            375                 380                 385

atc tgg act agt gaa ttc atc ctt gcg tgc cag caa atg act ata gct     1318
Ile Trp Thr Ser Glu Phe Ile Leu Ala Cys Gln Gln Met Thr Ile Ala
        390                 395                 400

ggg gca gtg gtt act tgt tat ttc aac aga agt aaa aat gat cct cct     1366
Gly Ala Val Val Thr Cys Tyr Phe Asn Arg Ser Lys Asn Asp Pro Pro
```

-continued

```
    405                 410                 415

gat cat ccc atc ctt tcg tct ctc tcc att ctc ttc ttc tac cat caa     1414
Asp His Pro Ile Leu Ser Ser Leu Ser Ile Leu Phe Phe Tyr His Gln
420                 425                 430                 435

gga acc att gtg aaa ggg tca ttt tta atc tct gtg gtg agg att ccg     1462
Gly Thr Ile Val Lys Gly Ser Phe Leu Ile Ser Val Val Arg Ile Pro
                440                 445                 450

aga atc att gtc atg tac atg caa aac gca ctg aaa gaa cag cag cat     1510
Arg Ile Ile Val Met Tyr Met Gln Asn Ala Leu Lys Glu Gln Gln His
            455                 460                 465

ggt gca ttg tcc agg tac ctg ttc cga tgc tgc tac tgc tgt ttc tgg     1558
Gly Ala Leu Ser Arg Tyr Leu Phe Arg Cys Cys Tyr Cys Cys Phe Trp
        470                 475                 480

tgt ctt gac aaa tac ctg ctc cat ctc aac cag aat gca tat act aca     1606
Cys Leu Asp Lys Tyr Leu Leu His Leu Asn Gln Asn Ala Tyr Thr Thr
    485                 490                 495

act gct att aat ggg aca gat ttc tgt aca tca gca aaa gat gca ttc     1654
Thr Ala Ile Asn Gly Thr Asp Phe Cys Thr Ser Ala Lys Asp Ala Phe
500                 505                 510                 515

aaa atc ttg tcc aag aac tca agt cac ttt aca tct att aac tgc ttt     1702
Lys Ile Leu Ser Lys Asn Ser Ser His Phe Thr Ser Ile Asn Cys Phe
                520                 525                 530

gga gac ttc ata att ttt cta gga aag gtg tta gtg gtg tgt ttc act     1750
Gly Asp Phe Ile Ile Phe Leu Gly Lys Val Leu Val Val Cys Phe Thr
            535                 540                 545

gtt ttt gga gga ctc atg gct ttt aac tac aat cgg gca ttc cag gtg     1798
Val Phe Gly Gly Leu Met Ala Phe Asn Tyr Asn Arg Ala Phe Gln Val
        550                 555                 560

tgg gca gtc cct ctg tta ttg gta gct ttt ttt gcc tac tta gta gcc     1846
Trp Ala Val Pro Leu Leu Leu Val Ala Phe Phe Ala Tyr Leu Val Ala
    565                 570                 575

cat agt ttt tta tct gtg ttt gaa act gtg ctg gat gca ctt ttc ctg     1894
His Ser Phe Leu Ser Val Phe Glu Thr Val Leu Asp Ala Leu Phe Leu
580                 585                 590                 595

tgt ttt gct gtt gat ctg gaa aca aat gat gga tcg tca gaa aag ccc     1942
Cys Phe Ala Val Asp Leu Glu Thr Asn Asp Gly Ser Ser Glu Lys Pro
                600                 605                 610

tac ttt atg gat caa gaa ttt ctg agt ttc gta aaa agg agc aac aaa     1990
Tyr Phe Met Asp Gln Glu Phe Leu Ser Phe Val Lys Arg Ser Asn Lys
            615                 620                 625

tta aac aat gca agg gca cag cag gac aag cac tca tta agg aat gag     2038
Leu Asn Asn Ala Arg Ala Gln Gln Asp Lys His Ser Leu Arg Asn Glu
        630                 635                 640

gag gga aca gaa ctc cag gcc att gtg aga tag atacccattt aggtatctgt   2091
Glu Gly Thr Glu Leu Gln Ala Ile Val Arg  *
    645                 650

acctggaaaa catttccttc taagagccat ttacagaata gaagatgaga ccactagaga   2151

aaagttagtg aatttttttt taaaagacct aataaaccct attcttcctc aaaaaaaaaa   2211

aaaaaaaaaa aaaaaaaaaa aa                                            2233


<210> SEQ ID NO 73
<211> LENGTH: 653
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 73
```

-continued

```
Met His Cys Leu Gly Ala Glu Tyr Leu Val Ser Ala Glu Gly Ala Pro
 1               5                  10                  15

Arg Gln Arg Glu Trp Arg Pro Gln Ile Tyr Arg Lys Cys Thr Asp Thr
            20                  25                  30

Ala Trp Leu Phe Leu Phe Phe Leu Phe Trp Thr Gly Leu Val Phe Ile
        35                  40                  45

Met Gly Tyr Ser Val Val Ala Gly Ala Ala Gly Arg Leu Leu Phe Gly
    50                  55                  60

Tyr Asp Ser Phe Gly Asn Met Cys Gly Lys Lys Asn Ser Pro Val Glu
65                  70                  75                  80

Gly Ala Pro Leu Ser Gly Gln Asp Met Thr Leu Lys Lys His Val Phe
                85                  90                  95

Phe Met Asn Ser Cys Asn Leu Glu Val Lys Gly Thr Gln Leu Asn Arg
            100                 105                 110

Met Ala Leu Cys Val Ser Asn Cys Pro Glu Glu Gln Leu Asp Ser Leu
        115                 120                 125

Glu Glu Val Gln Phe Phe Ala Asn Thr Ser Gly Ser Phe Leu Cys Val
    130                 135                 140

Tyr Ser Leu Asn Ser Phe Asn Tyr Thr His Ser Pro Lys Ala Asp Ser
145                 150                 155                 160

Leu Cys Pro Arg Leu Pro Val Pro Pro Ser Lys Ser Phe Pro Leu Phe
                165                 170                 175

Asn Arg Cys Val Pro Gln Thr Pro Glu Cys Tyr Ser Leu Phe Ala Ser
            180                 185                 190

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

Met Ser Gly Arg Asp Thr Ile Leu Gly Leu Cys Ile Leu Ala Leu Ala
    210                 215                 220

Leu Ser Leu Ala Met Met Phe Thr Phe Arg Phe Ile Thr Thr Leu Leu
225                 230                 235                 240

Val His Ile Phe Ile Ser Leu Val Ile Leu Gly Leu Leu Phe Val Cys
                245                 250                 255

Gly Val Leu Trp Trp Leu Tyr Tyr Asp Tyr Thr Asn Asp Leu Ser Ile
            260                 265                 270

Glu Leu Asp Thr Glu Arg Glu Asn Met Lys Cys Val Leu Gly Phe Ala
        275                 280                 285

Ile Val Ser Thr Gly Ile Thr Ala Val Leu Leu Val Leu Ile Phe Val
    290                 295                 300

Leu Arg Lys Arg Ile Lys Leu Thr Val Glu Leu Phe Gln Ile Thr Asn
305                 310                 315                 320

Lys Ala Ile Ser Ser Ala Pro Phe Leu Leu Phe Gln Pro Leu Trp Thr
                325                 330                 335

Phe Ala Ile Leu Ile Phe Phe Trp Val Leu Trp Val Ala Val Leu Leu
            340                 345                 350

Ser Leu Gly Thr Ala Gly Ala Ala Gln Val Met Glu Gly Gly Gln Val
        355                 360                 365

Glu Tyr Lys Pro Leu Ser Gly Ile Arg Tyr Met Trp Ser Tyr His Leu
    370                 375                 380

Ile Gly Leu Ile Trp Thr Ser Glu Phe Ile Leu Ala Cys Gln Gln Met
385                 390                 395                 400

Thr Ile Ala Gly Ala Val Val Thr Cys Tyr Phe Asn Arg Ser Lys Asn
```

-continued

```
                405                 410                 415

Asp Pro Pro Asp His Pro Ile Leu Ser Ser Leu Ser Ile Leu Phe Phe
            420                 425                 430

Tyr His Gln Gly Thr Ile Val Lys Gly Ser Phe Leu Ile Ser Val Val
        435                 440                 445

Arg Ile Pro Arg Ile Ile Val Met Tyr Met Gln Asn Ala Leu Lys Glu
    450                 455                 460

Gln Gln His Gly Ala Leu Ser Arg Tyr Leu Phe Arg Cys Cys Tyr Cys
465                 470                 475                 480

Cys Phe Trp Cys Leu Asp Lys Tyr Leu Leu His Leu Asn Gln Asn Ala
                485                 490                 495

Tyr Thr Thr Thr Ala Ile Asn Gly Thr Asp Phe Cys Thr Ser Ala Lys
            500                 505                 510

Asp Ala Phe Lys Ile Leu Ser Lys Asn Ser Ser His Phe Thr Ser Ile
        515                 520                 525

Asn Cys Phe Gly Asp Phe Ile Ile Phe Leu Gly Lys Val Leu Val Val
    530                 535                 540

Cys Phe Thr Val Phe Gly Gly Leu Met Ala Phe Asn Tyr Asn Arg Ala
545                 550                 555                 560

Phe Gln Val Trp Ala Val Pro Leu Leu Leu Val Ala Phe Phe Ala Tyr
                565                 570                 575

Leu Val Ala His Ser Phe Leu Ser Val Phe Glu Thr Val Leu Asp Ala
            580                 585                 590

Leu Phe Leu Cys Phe Ala Val Asp Leu Glu Thr Asn Asp Gly Ser Ser
        595                 600                 605

Glu Lys Pro Tyr Phe Met Asp Gln Glu Phe Leu Ser Phe Val Lys Arg
    610                 615                 620

Ser Asn Lys Leu Asn Asn Ala Arg Ala Gln Gln Asp Lys His Ser Leu
625                 630                 635                 640

Arg Asn Glu Glu Gly Thr Glu Leu Gln Ala Ile Val Arg
                645                 650


<210> SEQ ID NO 74
<211> LENGTH: 1962
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(1962)

<400> SEQUENCE: 74

atg cac tgc ctg ggc gcc gag tac ctg gtt tct gca gaa gga gcc cct       48
Met His Cys Leu Gly Ala Glu Tyr Leu Val Ser Ala Glu Gly Ala Pro
 1               5                  10                  15

agg caa agg gag tgg cga ccc cag att tat agg aaa tgc aca gat acg       96
Arg Gln Arg Glu Trp Arg Pro Gln Ile Tyr Arg Lys Cys Thr Asp Thr
            20                  25                  30

gca tgg tta ttc ctg ttc ttt ctc ttt tgg act ggt ttg gtg ttt atc      144
Ala Trp Leu Phe Leu Phe Phe Leu Phe Trp Thr Gly Leu Val Phe Ile
        35                  40                  45

atg ggc tac tcg gtg gtg gct gga gcc gcg gga aga ctc ctc ttt ggc      192
Met Gly Tyr Ser Val Val Ala Gly Ala Ala Gly Arg Leu Leu Phe Gly
    50                  55                  60

tat gac agc ttt ggc aac atg tgt ggc aag aag aac tcc ccc gtg gaa      240
Tyr Asp Ser Phe Gly Asn Met Cys Gly Lys Lys Asn Ser Pro Val Glu
```

-continued

```
 65                  70                  75                  80

ggg gcc cct ctt tca ggg cag gac atg acc cta aaa aaa cac gtg ttc      288
Gly Ala Pro Leu Ser Gly Gln Asp Met Thr Leu Lys Lys His Val Phe
                 85                  90                  95

ttt atg aat tcc tgc aac ctg gaa gtc aaa ggt acg cag ctc aac cgc      336
Phe Met Asn Ser Cys Asn Leu Glu Val Lys Gly Thr Gln Leu Asn Arg
            100                 105                 110

atg gcc ctc tgt gta tcc aac tgc cct gaa gag cag ctt gac tcc ctg      384
Met Ala Leu Cys Val Ser Asn Cys Pro Glu Glu Gln Leu Asp Ser Leu
        115                 120                 125

gaa gag gtc cag ttc ttt gca aac acc agt ggg tcc ttc ctg tgt gtt      432
Glu Glu Val Gln Phe Phe Ala Asn Thr Ser Gly Ser Phe Leu Cys Val
    130                 135                 140

tat agt ttg aat tcc ttc aac tat acc cac agt cca aaa gca gac tca      480
Tyr Ser Leu Asn Ser Phe Asn Tyr Thr His Ser Pro Lys Ala Asp Ser
145                 150                 155                 160

ctg tgt ccc agg cta cca gtt cct cca agc aag tca ttt ccc tta ttt      528
Leu Cys Pro Arg Leu Pro Val Pro Pro Ser Lys Ser Phe Pro Leu Phe
                165                 170                 175

aac cga tgt gtc cct caa aca cct gag tgc tac tcc cta ttt gca tct      576
Asn Arg Cys Val Pro Gln Thr Pro Glu Cys Tyr Ser Leu Phe Ala Ser
            180                 185                 190

gtt ttg ata aat gat gtt gac acc ctc cac cga att cta agt gga atc      624
Val Leu Ile Asn Asp Val Asp Thr Leu His Arg Ile Leu Ser Gly Ile
        195                 200                 205

atg tcg gga aga gat aca atc ctt ggc ctg tgt atc ctc gca tta gcc      672
Met Ser Gly Arg Asp Thr Ile Leu Gly Leu Cys Ile Leu Ala Leu Ala
    210                 215                 220

ttg tct ttg gcc atg atg ttt acc ttc aga ttc atc acc acc ctt ctg      720
Leu Ser Leu Ala Met Met Phe Thr Phe Arg Phe Ile Thr Thr Leu Leu
225                 230                 235                 240

gtt cac att ttc att tca ttg gtt att ttg gga ttg ttg ttt gtc tgc      768
Val His Ile Phe Ile Ser Leu Val Ile Leu Gly Leu Leu Phe Val Cys
                245                 250                 255

ggt gtt tta tgg tgg ctg tat tat gac tat acc aac gac ctc agc ata      816
Gly Val Leu Trp Trp Leu Tyr Tyr Asp Tyr Thr Asn Asp Leu Ser Ile
            260                 265                 270

gaa ttg gac aca gaa agg gaa aat atg aag tgc gtg ctg ggg ttt gct      864
Glu Leu Asp Thr Glu Arg Glu Asn Met Lys Cys Val Leu Gly Phe Ala
        275                 280                 285

atc gta tcc aca ggc atc acg gca gtg ctg ctc gtc ttg att ttt gtt      912
Ile Val Ser Thr Gly Ile Thr Ala Val Leu Leu Val Leu Ile Phe Val
    290                 295                 300

ctc aga aag aga ata aaa ttg aca gtt gag ctt ttc caa atc aca aat      960
Leu Arg Lys Arg Ile Lys Leu Thr Val Glu Leu Phe Gln Ile Thr Asn
305                 310                 315                 320

aaa gcc atc agc agt gct ccc ttc ctg ctg ttc cag cca ctg tgg aca     1008
Lys Ala Ile Ser Ser Ala Pro Phe Leu Leu Phe Gln Pro Leu Trp Thr
                325                 330                 335

ttt gcc atc ctc att ttc ttc tgg gtc ctc tgg gtg gct gtg ctg ctg     1056
Phe Ala Ile Leu Ile Phe Phe Trp Val Leu Trp Val Ala Val Leu Leu
            340                 345                 350

agc ctg gga act gca gga gct gcc cag gtt atg gaa ggc ggc caa gtg     1104
Ser Leu Gly Thr Ala Gly Ala Ala Gln Val Met Glu Gly Gly Gln Val
        355                 360                 365

gaa tat aag ccc ctt tcg ggc att cgg tac atg tgg tcg tac cat tta     1152
Glu Tyr Lys Pro Leu Ser Gly Ile Arg Tyr Met Trp Ser Tyr His Leu
```

-continued

```
    370                 375                 380

att ggc ctc atc tgg act agt gaa ttc atc ctt gcg tgc cag caa atg     1200
Ile Gly Leu Ile Trp Thr Ser Glu Phe Ile Leu Ala Cys Gln Gln Met
385                 390                 395                 400

act ata gct ggg gca gtg gtt act tgt tat ttc aac aga agt aaa aat     1248
Thr Ile Ala Gly Ala Val Val Thr Cys Tyr Phe Asn Arg Ser Lys Asn
                405                 410                 415

gat cct cct gat cat ccc atc ctt tcg tct ctc tcc att ctc ttc ttc     1296
Asp Pro Pro Asp His Pro Ile Leu Ser Ser Leu Ser Ile Leu Phe Phe
            420                 425                 430

tac cat caa gga acc att gtg aaa ggg tca ttt tta atc tct gtg gtg     1344
Tyr His Gln Gly Thr Ile Val Lys Gly Ser Phe Leu Ile Ser Val Val
        435                 440                 445

agg att ccg aga atc att gtc atg tac atg caa aac gca ctg aaa gaa     1392
Arg Ile Pro Arg Ile Ile Val Met Tyr Met Gln Asn Ala Leu Lys Glu
    450                 455                 460

cag cag cat ggt gca ttg tcc agg tac ctg ttc cga tgc tgc tac tgc     1440
Gln Gln His Gly Ala Leu Ser Arg Tyr Leu Phe Arg Cys Cys Tyr Cys
465                 470                 475                 480

tgt ttc tgg tgt ctt gac aaa tac ctg ctc cat ctc aac cag aat gca     1488
Cys Phe Trp Cys Leu Asp Lys Tyr Leu Leu His Leu Asn Gln Asn Ala
                485                 490                 495

tat act aca act gct att aat ggg aca gat ttc tgt aca tca gca aaa     1536
Tyr Thr Thr Thr Ala Ile Asn Gly Thr Asp Phe Cys Thr Ser Ala Lys
            500                 505                 510

gat gca ttc aaa atc ttg tcc aag aac tca agt cac ttt aca tct att     1584
Asp Ala Phe Lys Ile Leu Ser Lys Asn Ser Ser His Phe Thr Ser Ile
        515                 520                 525

aac tgc ttt gga gac ttc ata att ttt cta gga aag gtg tta gtg gtg     1632
Asn Cys Phe Gly Asp Phe Ile Ile Phe Leu Gly Lys Val Leu Val Val
    530                 535                 540

tgt ttc act gtt ttt gga gga ctc atg gct ttt aac tac aat cgg gca     1680
Cys Phe Thr Val Phe Gly Gly Leu Met Ala Phe Asn Tyr Asn Arg Ala
545                 550                 555                 560

ttc cag gtg tgg gca gtc cct ctg tta ttg gta gct ttt ttt gcc tac     1728
Phe Gln Val Trp Ala Val Pro Leu Leu Leu Val Ala Phe Phe Ala Tyr
                565                 570                 575

tta gta gcc cat agt ttt tta tct gtg ttt gaa act gtg ctg gat gca     1776
Leu Val Ala His Ser Phe Leu Ser Val Phe Glu Thr Val Leu Asp Ala
            580                 585                 590

ctt ttc ctg tgt ttt gct gtt gat ctg gaa aca aat gat gga tcg tca     1824
Leu Phe Leu Cys Phe Ala Val Asp Leu Glu Thr Asn Asp Gly Ser Ser
        595                 600                 605

gaa aag ccc tac ttt atg gat caa gaa ttt ctg agt ttc gta aaa agg     1872
Glu Lys Pro Tyr Phe Met Asp Gln Glu Phe Leu Ser Phe Val Lys Arg
    610                 615                 620

agc aac aaa tta aac aat gca agg gca cag cag gac aag cac tca tta     1920
Ser Asn Lys Leu Asn Asn Ala Arg Ala Gln Gln Asp Lys His Ser Leu
625                 630                 635                 640

agg aat gag gag gga aca gaa ctc cag gcc att gtg aga tag             1962
Arg Asn Glu Glu Gly Thr Glu Leu Gln Ala Ile Val Arg  *
                645                 650
```

<210> SEQ ID NO 75
<211> LENGTH: 654
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

-continued

<400> SEQUENCE: 75

Met Gly Cys Cys Ser Ser Ala Ser Ser Ala Ala Gln Ser Ser Lys Arg
1 5 10 15

Glu Trp Lys Pro Leu Glu Asp Arg Ser Cys Thr Asp Ile Pro Trp Leu
20 25 30

Leu Leu Phe Ile Leu Phe Cys Ile Gly Met Gly Phe Ile Cys Gly Phe
35 40 45

Ser Ile Ala Thr Gly Ala Ala Ala Arg Leu Val Ser Gly Tyr Asp Ser
50 55 60

Tyr Gly Asn Ile Cys Gly Gln Lys Asn Thr Lys Leu Glu Ala Ile Pro
65 70 75 80

Asn Ser Gly Met Asp His Thr Gln Arg Lys Tyr Val Phe Phe Leu Asp
85 90 95

Pro Cys Asn Leu Asp Leu Ile Asn Arg Lys Ile Lys Ser Val Ala Leu
100 105 110

Cys Val Ala Ala Cys Pro Arg Gln Glu Leu Lys Thr Leu Ser Asp Val
115 120 125

Gln Lys Phe Ala Glu Ile Asn Gly Ser Ala Leu Cys Ser Tyr Asn Leu
130 135 140

Lys Pro Ser Glu Tyr Thr Thr Ser Pro Lys Ser Ser Val Leu Cys Pro
145 150 155 160

Lys Leu Pro Val Pro Ala Ser Ala Pro Ile Pro Phe Phe His Arg Cys
165 170 175

Ala Pro Val Asn Ile Ser Cys Tyr Ala Lys Phe Ala Glu Ala Leu Ile
180 185 190

Thr Phe Val Ser Asp Asn Ser Val Leu His Arg Leu Ile Ser Gly Val
195 200 205

Met Thr Ser Lys Glu Ile Ile Leu Gly Leu Cys Leu Leu Ser Leu Val
210 215 220

Leu Ser Met Ile Leu Met Val Ile Ile Arg Tyr Ile Ser Arg Val Leu
225 230 235 240

Val Trp Ile Leu Thr Ile Leu Val Ile Leu Gly Ser Leu Gly Gly Thr
245 250 255

Gly Val Leu Trp Trp Leu Tyr Ala Lys Gln Arg Arg Ser Pro Lys Glu
260 265 270

Thr Val Thr Pro Glu Gln Leu Gln Ile Ala Glu Asp Asn Leu Arg Ala
275 280 285

Leu Leu Ile Tyr Ala Ile Ser Ala Thr Val Phe Thr Val Ile Leu Phe
290 295 300

Leu Ile Met Leu Val Met Arg Lys Arg Val Ala Leu Thr Ile Ala Leu
305 310 315 320

Phe His Val Ala Gly Lys Val Phe Ile His Leu Pro Leu Leu Val Phe
325 330 335

Gln Pro Phe Trp Thr Phe Phe Ala Leu Val Leu Phe Trp Val Tyr Trp
340 345 350

Ile Met Thr Leu Leu Phe Leu Gly Thr Thr Gly Ser Pro Val Gln Asn
355 360 365

Glu Gln Gly Phe Val Glu Phe Lys Ile Ser Gly Pro Leu Gln Tyr Met
370 375 380

Trp Trp Tyr His Val Val Gly Leu Ile Trp Ile Ser Glu Phe Ile Leu
385 390 395 400

-continued

```
Ala Cys Gln Gln Met Thr Val Ala Gly Ala Val Val Thr Tyr Tyr Phe
                405                 410                 415

Thr Arg Asp Lys Arg Asn Leu Pro Phe Thr Pro Ile Leu Ala Ser Val
            420                 425                 430

Asn Arg Leu Ile Arg Tyr His Leu Gly Thr Val Ala Lys Gly Ser Phe
        435                 440                 445

Ile Ile Thr Leu Val Lys Ile Pro Arg Met Ile Leu Met Tyr Ile His
    450                 455                 460

Ser Gln Leu Lys Gly Lys Glu Asn Ala Cys Ala Arg Cys Val Leu Lys
465                 470                 475                 480

Ser Cys Ile Cys Cys Leu Trp Cys Leu Glu Lys Cys Leu Asn Tyr Leu
                485                 490                 495

Asn Gln Asn Ala Tyr Thr Ala Thr Ala Ile Asn Ser Thr Asn Phe Cys
            500                 505                 510

Thr Ser Ala Lys Asp Ala Phe Val Ile Leu Val Glu Asn Ala Leu Arg
        515                 520                 525

Val Ala Thr Ile Asn Thr Val Gly Asp Phe Met Leu Phe Leu Gly Lys
    530                 535                 540

Val Leu Ile Val Cys Ser Thr Gly Leu Ala Gly Ile Met Leu Leu Asn
545                 550                 555                 560

Tyr Gln Gln Asp Tyr Thr Val Trp Val Leu Pro Leu Ile Ile Val Cys
                565                 570                 575

Leu Phe Ala Phe Leu Val Ala His Cys Phe Leu Ser Ile Tyr Glu Met
            580                 585                 590

Val Val Asp Val Leu Phe Leu Cys Phe Ala Ile Asp Thr Lys Tyr Asn
        595                 600                 605

Asp Gly Ser Pro Gly Arg Glu Phe Tyr Met Asp Lys Val Leu Met Glu
    610                 615                 620

Phe Val Glu Asn Ser Arg Lys Ala Met Lys Glu Ala Gly Lys Gly Gly
625                 630                 635                 640

Val Ala Asp Ser Arg Glu Leu Lys Pro Met Leu Lys Lys Arg
                645                 650


<210> SEQ ID NO 76
<211> LENGTH: 706
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 76

Met Gly Asp Glu Arg Pro His Tyr Tyr Gly Lys His Gly Thr Pro Gln
 1               5                  10                  15

Lys Tyr Asp Pro Thr Phe Lys Gly Pro Ile Tyr Asn Arg Gly Cys Thr
            20                  25                  30

Asp Ile Ile Cys Cys Val Phe Leu Leu Leu Ala Ile Val Gly Tyr Val
        35                  40                  45

Ala Val Gly Ile Ile Ala Trp Thr His Gly Asp Pro Arg Lys Val Ile
    50                  55                  60

Tyr Pro Thr Asp Ser Arg Gly Glu Phe Cys Gly Gln Lys Gly Thr Lys
65                  70                  75                  80

Asn Glu Asn Lys Pro Tyr Leu Phe Tyr Phe Asn Ile Val Lys Cys Ala
                85                  90                  95

Ser Pro Leu Val Leu Leu Glu Phe Gln Cys Pro Thr Pro Gln Ile Cys
```

-continued

```
            100                 105                 110

Val Glu Lys Cys Pro Asp Arg Tyr Leu Thr Tyr Leu Asn Ala Arg Ser
        115                 120                 125

Ser Arg Asp Phe Glu Tyr Tyr Lys Gln Phe Cys Val Pro Gly Phe Lys
    130                 135                 140

Asn Asn Lys Gly Val Ala Glu Val Leu Arg Asp Gly Asp Cys Pro Ala
145                 150                 155                 160

Val Leu Ile Pro Ser Lys Pro Leu Ala Arg Arg Cys Phe Pro Ala Ile
                165                 170                 175

His Ala Tyr Lys Gly Val Leu Met Val Gly Asn Glu Thr Thr Tyr Glu
            180                 185                 190

Asp Gly His Gly Ser Arg Lys Asn Ile Thr Asp Leu Val Glu Gly Ala
        195                 200                 205

Lys Lys Ala Asn Gly Val Leu Glu Ala Arg Gln Leu Ala Met Arg Ile
    210                 215                 220

Phe Glu Asp Tyr Thr Val Ser Trp Tyr Trp Ile Ile Ile Gly Leu Val
225                 230                 235                 240

Ile Ala Met Ala Met Ser Leu Leu Phe Ile Ile Leu Leu Arg Phe Leu
                245                 250                 255

Ala Gly Ile Met Val Trp Val Met Ile Ile Met Val Ile Leu Val Leu
            260                 265                 270

Gly Tyr Gly Ile Phe His Cys Tyr Met Glu Tyr Ser Arg Leu Arg Gly
        275                 280                 285

Glu Ala Gly Ser Asp Val Ser Leu Val Asp Leu Gly Phe Gln Thr Asp
    290                 295                 300

Phe Arg Val Tyr Leu His Leu Arg Gln Thr Trp Leu Ala Phe Met Ile
305                 310                 315                 320

Ile Leu Ser Ile Leu Glu Val Ile Ile Ile Leu Leu Leu Ile Phe Leu
                325                 330                 335

Arg Lys Arg Ile Leu Ile Ala Ile Ala Leu Ile Lys Glu Ala Ser Arg
            340                 345                 350

Ala Val Gly Tyr Val Met Cys Ser Leu Leu Tyr Pro Leu Val Thr Phe
        355                 360                 365

Phe Leu Leu Cys Leu Cys Ile Ala Tyr Trp Ala Ser Thr Ala Val Phe
    370                 375                 380

Leu Ser Thr Ser Asn Glu Ala Val Tyr Lys Ile Phe Asp Asp Ser Pro
385                 390                 395                 400

Cys Pro Phe Thr Ala Lys Thr Cys Asn Pro Glu Thr Phe Pro Ser Ser
                405                 410                 415

Asn Glu Ser Arg Gln Cys Pro Asn Ala Arg Cys Gln Phe Ala Phe Tyr
            420                 425                 430

Gly Gly Glu Ser Gly Tyr His Arg Ala Leu Leu Gly Leu Gln Ile Phe
        435                 440                 445

Asn Ala Phe Met Phe Phe Trp Leu Ala Asn Phe Val Leu Ala Leu Gly
    450                 455                 460

Gln Val Thr Leu Ala Gly Ala Phe Ala Ser Tyr Tyr Trp Ala Leu Arg
465                 470                 475                 480

Lys Pro Asp Asp Leu Pro Ala Phe Pro Leu Phe Ser Ala Phe Gly Arg
                485                 490                 495

Ala Leu Arg Tyr His Thr Gly Ser Leu Ala Phe Gly Ala Leu Ile Leu
            500                 505                 510
```

-continued

```
Ala Ile Val Gln Ile Ile Arg Val Ile Leu Glu Tyr Leu Asp Gln Arg
        515                 520                 525

Leu Lys Gly Ala Glu Asn Lys Phe Ala Lys Cys Leu Met Thr Cys Leu
    530                 535                 540

Lys Cys Cys Phe Trp Cys Leu Glu Lys Phe Ile Lys Phe Leu Asn Arg
545                 550                 555                 560

Asn Ala Tyr Ile Met Ile Ala Ile Tyr Gly Thr Asn Phe Cys Thr Ser
                565                 570                 575

Ala Arg Asn Ala Phe Phe Leu Leu Met Arg Asn Ile Ile Arg Val Ala
            580                 585                 590

Val Leu Asp Lys Val Thr Asp Phe Leu Phe Leu Leu Gly Lys Leu Leu
        595                 600                 605

Ile Val Gly Ser Val Gly Ile Leu Ala Phe Phe Phe Phe Thr His Arg
    610                 615                 620

Ile Arg Ile Val Gln Asp Thr Ala Pro Pro Leu Asn Tyr Tyr Trp Val
625                 630                 635                 640

Pro Ile Leu Thr Val Ile Val Gly Ser Tyr Leu Ile Ala His Gly Phe
                645                 650                 655

Phe Ser Val Tyr Gly Met Cys Val Asp Thr Leu Phe Leu Cys Phe Leu
            660                 665                 670

Glu Asp Leu Glu Arg Asn Asp Gly Ser Ala Glu Arg Pro Tyr Phe Met
        675                 680                 685

Ser Ser Thr Leu Lys Lys Leu Leu Asn Lys Thr Asn Lys Lys Ala Ala
    690                 695                 700

Glu Ser
705


<210> SEQ ID NO 77
<211> LENGTH: 653
<212> TYPE: PRT
<213> ORGANISM: Rattus norvegicus

<400> SEQUENCE: 77

Met Gly Cys Cys Ser Ser Ala Ser Ala Ala Gln Ser Ser Lys Arg Glu
 1               5                  10                  15

Trp Lys Pro Leu Glu Asp Arg Ser Cys Thr Asp Ile Pro Trp Leu Leu
            20                  25                  30

Leu Phe Val Leu Phe Cys Ile Gly Met Gly Phe Ile Cys Gly Phe Ser
        35                  40                  45

Val Ala Thr Gly Ala Ala Ala Arg Leu Val Ser Gly Tyr Asp Ser Tyr
    50                  55                  60

Gly Asn Ile Cys Gly Gln Arg Asn Ala Lys Leu Glu Ala Ile Ala Asn
65                  70                  75                  80

Ser Gly Leu Asp His Thr His Arg Lys Tyr Val Phe Phe Leu Asp Pro
                85                  90                  95

Cys Asn Leu Asp Leu Ile Asn Arg Lys Ile Lys Ser Met Ala Leu Cys
            100                 105                 110

Val Ala Ala Cys Pro Arg Gln Glu Leu Lys Thr Leu Ser Asp Val Gln
        115                 120                 125

Lys Phe Ala Glu Ile Asn Gly Ser Ala Leu Cys Ser Tyr Asn Ile Lys
    130                 135                 140

Pro Ser Glu Tyr Thr Leu Thr Ala Lys Ser Ser Ala Phe Cys Pro Lys
```

-continued

```
145                 150                 155                 160

Leu Pro Val Pro Ala Ser Ala Pro Ile Pro Phe Phe His Arg Cys Ala
                165                 170                 175

Pro Val Asn Ile Ser Cys Tyr Ala Lys Phe Ala Glu Ala Leu Ile Thr
            180                 185                 190

Phe Val Ser Asp Asn Ser Val Leu His Arg Leu Ile Ser Gly Val Met
        195                 200                 205

Thr Ser Lys Glu Ile Ile Leu Gly Leu Cys Leu Leu Ser Leu Val Leu
    210                 215                 220

Ser Met Ile Leu Met Val Ile Ile Arg Tyr Ile Ser Arg Val Leu Val
225                 230                 235                 240

Trp Ile Leu Thr Ile Leu Val Ile Leu Gly Ser Leu Gly Gly Thr Gly
                245                 250                 255

Val Leu Trp Trp Leu Tyr Ala Lys Gln Arg Ser Ser Pro Lys Glu Thr
            260                 265                 270

Val Ile Pro Glu Gln Leu Gln Ile Ala Glu Asp Asn Leu Arg Ala Leu
        275                 280                 285

Leu Ile Tyr Ala Ile Ser Ala Thr Val Phe Thr Val Ile Leu Phe Leu
    290                 295                 300

Ile Met Leu Val Met Arg Lys Arg Val Ala Leu Thr Ile Ala Leu Phe
305                 310                 315                 320

His Val Ala Gly Lys Val Phe Ile His Leu Pro Leu Leu Val Phe Gln
                325                 330                 335

Pro Phe Trp Thr Phe Phe Ala Leu Val Leu Phe Trp Ala Tyr Trp Ile
            340                 345                 350

Met Thr Leu Leu Phe Leu Gly Thr Thr Gly Ser Ala Val Gln Asn Glu
        355                 360                 365

Gln Gly Phe Val Glu Tyr Lys Ile Ser Gly Pro Leu Gln Tyr Met Trp
    370                 375                 380

Trp Tyr His Val Val Gly Leu Ile Trp Ile Ser Glu Phe Ile Leu Ala
385                 390                 395                 400

Cys Gln Gln Met Thr Val Ala Gly Ala Val Val Thr Tyr Tyr Phe Thr
                405                 410                 415

Arg Asp Lys Arg Asn Leu Pro Phe Thr Pro Ile Leu Ala Ser Val Asn
            420                 425                 430

Arg Leu Ile Arg Tyr His Leu Gly Thr Val Ala Lys Gly Ser Phe Ile
        435                 440                 445

Ile Thr Leu Val Lys Ile Pro Arg Met Ile Leu Met Tyr Ile His Ser
    450                 455                 460

Gln Leu Lys Gly Lys Glu Asn Ala Cys Ala Arg Cys Met Leu Lys Ser
465                 470                 475                 480

Cys Ile Cys Cys Leu Trp Cys Leu Glu Lys Cys Leu Ser Tyr Leu Asn
                485                 490                 495

Gln Asn Ala Tyr Thr Ala Thr Ala Ile Asn Ser Thr Asn Phe Cys Thr
            500                 505                 510

Ser Ala Lys Asp Ala Phe Val Ile Leu Val Glu Asn Ala Leu Arg Val
        515                 520                 525

Ala Ala Ile Asn Thr Val Gly Asp Phe Met Leu Phe Leu Gly Lys Val
    530                 535                 540

Leu Ile Val Cys Ser Thr Gly Leu Ala Gly Ile Met Leu Leu Asn Tyr
545                 550                 555                 560
```

-continued

```
Gln Gln Asp Tyr Thr Val Trp Val Leu Pro Leu Ile Ile Val Cys Leu
                565                 570                 575

Phe Ala Phe Leu Val Ala His Cys Phe Leu Ser Ile Tyr Glu Met Val
            580                 585                 590

Val Asp Val Leu Phe Leu Cys Phe Ala Ile Asp Thr Lys Tyr Asn Asp
        595                 600                 605

Gly Ser Pro Gly Arg Glu Phe Tyr Met Asp Lys Val Leu Met Glu Phe
    610                 615                 620

Val Glu Asn Ser Arg Lys Ala Met Lys Glu Ala Gly Lys Gly Gly Ala
625                 630                 635                 640

Ala Asp Ala Arg Glu Leu Lys Pro Met Leu Arg Lys Arg
                645                 650


<210> SEQ ID NO 78
<211> LENGTH: 646
<212> TYPE: PRT
<213> ORGANISM: Torpedo marmorata

<400> SEQUENCE: 78

Met Gly Cys Cys Gly Cys Gly Ser Glu Glu Gly Ser Val Arg Gln Trp
 1               5                  10                  15

Lys Pro Leu Glu Gln Arg Ser Cys Thr Asp Val Leu Trp Leu Leu Ile
            20                  25                  30

Phe Val Leu Phe Cys Ile Gly Met Ala Ile Ile Cys Gly Phe Ala Ile
        35                  40                  45

Ala Ser Gly Ala Ala Gln Arg Leu Val Phe Gly Tyr Asp Ser Tyr Gly
    50                  55                  60

Asn Ile Cys Gly His Lys Asn Thr Glu Ile Lys Asp Val Thr Met Ser
65                  70                  75                  80

Gly Leu Asp His Thr Asp Lys Lys Tyr Val Phe Phe Phe Glu Pro Cys
                85                  90                  95

Asn Trp Asp Met Val His Leu Lys Ile Leu Ser Val Ala Leu Cys Val
            100                 105                 110

Thr Lys Cys Pro Asp Met Asp Leu Lys Thr Leu Glu Asp Val Arg Asn
        115                 120                 125

Phe Ala Lys Tyr Asn Gly Ser Arg Leu Cys Leu Tyr Asn Leu Asp Pro
    130                 135                 140

Thr Gln Tyr Thr Ser Lys Asn Ser Lys Ser Cys Pro Ile Leu Pro Val
145                 150                 155                 160

Lys Ser Ser Lys Pro Ile Pro Phe Phe His Arg Cys Val Pro Met Asp
                165                 170                 175

Ser Gly Cys Lys Ile Asn Phe Lys Ala Leu Thr Thr Phe Val Ser Tyr
            180                 185                 190

Asn Ser Val Leu Gln Arg Val Ile Thr Gly Val Met Thr Ser Lys Glu
        195                 200                 205

Ile Ile Val Gly Leu Cys Leu Met Ser Leu Val Leu Ser Ile Leu Leu
    210                 215                 220

Met Val Ile Ile Arg Tyr Ile Ser Lys Val Leu Val Trp Ile Leu Ala
225                 230                 235                 240

Ile Leu Thr Ile Ile Gly Ser Ile Gly Gly Thr Ala Val Leu Trp Trp
                245                 250                 255

Leu Tyr Ala Asp His Lys Lys Thr Leu Lys Leu Asp Pro Ser Gln Gly
```

-continued

```
            260                 265                 270

Asp Val Ala Ala Asp Asn Val Thr Ala Leu Leu Val Cys Ala Ile Ile
        275                 280                 285

Ala Thr Val Ile Thr Val Ile Leu Leu Leu Leu Met Leu Ile Met Arg
    290                 295                 300

Lys Arg Val Ala Leu Thr Ile Ala Leu Phe His Val Ala Gly Lys Val
305                 310                 315                 320

Phe Ile His Ile Pro Phe Leu Ile Phe Gln Ser Leu Trp Thr Phe Leu
                325                 330                 335

Ala Leu Ala Phe Phe Trp Ile Tyr Trp Ile Ala Val Leu Leu Leu Leu
            340                 345                 350

Ala Thr Ala Gly Tyr Pro Gln Lys Lys Asp Gln Gly Tyr Val Glu Phe
        355                 360                 365

Lys Val Ser Gly Pro Leu Gln Tyr Thr Trp Ile Tyr His Leu Val Gly
    370                 375                 380

Leu Ile Trp Ile Ser Glu Phe Ile Leu Ala Cys Gln Gln Met Thr Ile
385                 390                 395                 400

Ala Gly Ala Val Val Thr Tyr Tyr Phe Thr Arg Asp Lys His Asn Leu
                405                 410                 415

Pro Ala Thr Pro Ile Leu Ala Ser Met Cys Arg Leu Ile Lys Tyr His
            420                 425                 430

Leu Gly Thr Val Ala Lys Gly Ser Phe Ile Ile Thr Leu Ile Lys Ile
        435                 440                 445

Pro Gln Met Ile Leu Val Tyr Ile His Ser Gln Leu Lys Gly Lys Glu
    450                 455                 460

Asn Ala Cys Ala Lys Cys Met Leu Lys Ala Cys Met Cys Cys Leu Trp
465                 470                 475                 480

Cys Leu Glu Lys Cys Leu Leu Tyr Leu Asn Arg Asn Ala Tyr Ile Ala
                485                 490                 495

Thr Ser Ile Asn Gly Thr Ser Phe Cys Thr Ser Ala Lys Asp Ala Ile
            500                 505                 510

Val Ile Leu Val Glu Asn Ala Met Arg Val Ala Ala Ile Asn Thr Val
        515                 520                 525

Gly Asp Phe Val Leu Phe Leu Gly Lys Leu Leu Ile Val Leu Val Thr
    530                 535                 540

Gly Phe Val Gly Ile Ile Leu Leu Asn Tyr Gln Arg Asp Tyr Thr Val
545                 550                 555                 560

Trp Val Leu Pro Leu Ile Ile Ile Cys Leu Phe Ala Phe Phe Val Ser
                565                 570                 575

His Cys Phe Leu Ser Ile Tyr Glu Met Val Val Asp Val Leu Phe Leu
            580                 585                 590

Cys Phe Ala Val Asp Cys Lys His Asn Asp Gly Ser Pro Gly Arg Glu
        595                 600                 605

Tyr Tyr Met Asp Lys Ser Leu Met Glu Phe Met Asp Glu Ser Arg Lys
    610                 615                 620

Ala Met Arg Ser Val Thr Gly Ser Gly Ala Glu Met Lys Ser Met Ala
625                 630                 635                 640

Ser Gly Ser Asp Asn Ala
                645
```

<210> SEQ ID NO 79

-continued

<211> LENGTH: 52
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (1)...(1)
<223> OTHER INFORMATION: Xaa at position 1 can be L, V or I
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (5)...(6)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (7)...(7)
<223> OTHER INFORMATION: Xaa at position 7 can be S or T
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (8)...(8)
<223> OTHER INFORMATION: Xaa at position 8 can be C or Y
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (10)...(10)
<223> OTHER INFORMATION: Xaa at position 10 can be F or W
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (11)...(13)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (15)...(15)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (17)...(18)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (20)...(20)
<223> OTHER INFORMATION: Xaa at position 20 can be L or I
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (21)...(25)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (26)...(26)
<223> OTHER INFORMATION: Xaa at position 26 can be I or R
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (27)...(29)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (32)...(32)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (34)...(37)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (39)...(40)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (41)...(42)
<223> OTHER INFORMATION: Xaa at positions 41-42 can be L or I
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (43)...(46)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (47)...(47)

-continued

<223> OTHER INFORMATION: Xaa at position 47 can be I or M
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (48)...(49)
<223> OTHER INFORMATION: Xaa = any amino acid.
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (50)...(50)
<223> OTHER INFORMATION: Xaa at position 50 can be V, M, or I
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (51)...(51)
<223> OTHER INFORMATION: Xaa at position 51 can be V or I
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (52)...(52)
<223> OTHER INFORMATION: Xaa at position 52 can be V or L

<400> SEQUENCE: 79

Xaa Ala Gly Ala Xaa Xaa Xaa Xaa Tyr Xaa Xaa Xaa Xaa Lys Xaa Pro
1 5 10 15

Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Tyr His Xaa
20 25 30

Gly Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa
50

<210> SEQ ID NO 80
<211> LENGTH: 49
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (3)...(3)
<223> OTHER INFORMATION: Xaa at position 3 can be E, R, or G
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (4)...(5)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (6)...(6)
<223> OTHER INFORMATION: Xaa at position 6 can be H or N
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (10)...(10)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (14)...(14)
<223> OTHER INFORMATION: Xaa at position 14 can be D or E
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (15)...(22)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (26)...(28)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (29)...(29)
<223> OTHER INFORMATION: Xaa at position 29 can be A or S
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (37)...(37)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:

-continued

<221> NAME/KEY: VARIANT
<222> LOCATION: (43)...(43)
<223> OTHER INFORMATION: X at position 43 can be F or I
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (44)...(44)
<223> OTHER INFORMATION: Xaa = any amino acid
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (45)...(45)
<223> OTHER INFORMATION: Xaa at position 45 can be I or L
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (47)...(48)
<223> OTHER INFORMATION: Xaa = any amino acid

<400> SEQUENCE: 80

Leu Lys Xaa Xaa Xaa Xaa Xaa Cys Cys Xaa Trp Cys Leu Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Asn Ala Tyr Xaa Xaa Xaa Xaa Ile Xaa Xaa
20 25 30

Xaa Xaa Phe Cys Xaa Ser Ala Lys Asp Ala Xaa Xaa Xaa Leu Xaa Xaa
35 40 45

Asn

<210> SEQ ID NO 81
<211> LENGTH: 4113
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (101)...(3691)

<400> SEQUENCE: 81

caaagagcgg ctcgcgcgct ggcctgggct ctaaccgagg agagatcccg ggagaactcc 60

agagctccgg gggagcgctc ctcggaagac cggggccaac atg cct gtg cgc agg 115
Met Pro Val Arg Arg
1 5

ggg cat gtg gca cca caa aat aca ttt ctg ggg acc atc att cgg aaa 163
Gly His Val Ala Pro Gln Asn Thr Phe Leu Gly Thr Ile Ile Arg Lys
10 15 20

ttt gaa ggg caa aat aaa aaa ttt atc att gca aat gcc aga gtg cag 211
Phe Glu Gly Gln Asn Lys Lys Phe Ile Ile Ala Asn Ala Arg Val Gln
25 30 35

aac tgt gcc atc att tat tgc aac gat ggg ttc tgt gag atg act ggt 259
Asn Cys Ala Ile Ile Tyr Cys Asn Asp Gly Phe Cys Glu Met Thr Gly
40 45 50

ttc tcc agg cca gat gtc atg caa aag cca tgc acc tgc gac ttt ctc 307
Phe Ser Arg Pro Asp Val Met Gln Lys Pro Cys Thr Cys Asp Phe Leu
55 60 65

cat gga ccc gag acc aag agg cat gat att gcc caa att gcc cag gca 355
His Gly Pro Glu Thr Lys Arg His Asp Ile Ala Gln Ile Ala Gln Ala
70 75 80 85

ttg ctg ggg tca gaa gag agg aaa gtg gag gtc acc tac tat cac aaa 403
Leu Leu Gly Ser Glu Glu Arg Lys Val Glu Val Thr Tyr Tyr His Lys
90 95 100

aat ggg tcc act ttt att tgt aac act cac ata att cca gtg aaa aac 451
Asn Gly Ser Thr Phe Ile Cys Asn Thr His Ile Ile Pro Val Lys Asn
105 110 115

caa gag ggc gtg gct atg atg ttc atc att aat ttt gaa tat gtg acg 499
Gln Glu Gly Val Ala Met Met Phe Ile Ile Asn Phe Glu Tyr Val Thr

-continued

```
        120                 125                 130

gat aat gaa aac gct gcc acc cca gag agg gta aac cca ata tta cca      547
Asp Asn Glu Asn Ala Ala Thr Pro Glu Arg Val Asn Pro Ile Leu Pro
    135                 140                 145

atc aaa act gta aac cgg aaa ttt ttt ggg ttc aaa ttc cct ggt ctg      595
Ile Lys Thr Val Asn Arg Lys Phe Phe Gly Phe Lys Phe Pro Gly Leu
150                 155                 160                 165

aga gtt ctc act tac aga aag cag tcc tta cca caa gaa gac ccc gat      643
Arg Val Leu Thr Tyr Arg Lys Gln Ser Leu Pro Gln Glu Asp Pro Asp
                170                 175                 180

gtg gtg gtc atc gat tca tct aaa cac agt gat gat tca gta gcc atg      691
Val Val Val Ile Asp Ser Ser Lys His Ser Asp Asp Ser Val Ala Met
            185                 190                 195

aag cat ttt aag tct cct aca aaa gaa agc tgc agc ccc tct gaa gca      739
Lys His Phe Lys Ser Pro Thr Lys Glu Ser Cys Ser Pro Ser Glu Ala
        200                 205                 210

gat gac aca aaa gct ttg ata cag ccc agc aaa tgt tct ccc ttg gtg      787
Asp Asp Thr Lys Ala Leu Ile Gln Pro Ser Lys Cys Ser Pro Leu Val
    215                 220                 225

aat ata tcc gga cct ctt gac cat tcc tct ccc aaa agg caa tgg gac      835
Asn Ile Ser Gly Pro Leu Asp His Ser Ser Pro Lys Arg Gln Trp Asp
230                 235                 240                 245

cga ctc tac cct gac atg ctg cag tca agt tcc cag ctg tcc cat tcc      883
Arg Leu Tyr Pro Asp Met Leu Gln Ser Ser Ser Gln Leu Ser His Ser
                250                 255                 260

aga tca agg gaa agc tta tgt agt ata cgg aga gca tct tcg gtc cat      931
Arg Ser Arg Glu Ser Leu Cys Ser Ile Arg Arg Ala Ser Ser Val His
            265                 270                 275

gat ata gaa gga ttc ggc gtc cac ccc aag aac ata ttt aga gac cga      979
Asp Ile Glu Gly Phe Gly Val His Pro Lys Asn Ile Phe Arg Asp Arg
        280                 285                 290

cat gcc agc gaa gac aat ggt cgc aat gtc aaa ggg cct ttt aat cat     1027
His Ala Ser Glu Asp Asn Gly Arg Asn Val Lys Gly Pro Phe Asn His
    295                 300                 305

atc aag tca agc ctc ctg gga tcc aca tca gat tca aac ctc aac aaa     1075
Ile Lys Ser Ser Leu Leu Gly Ser Thr Ser Asp Ser Asn Leu Asn Lys
310                 315                 320                 325

tac agc acc att aac aag att cca cag ctc act ctg aat ttt tca gag     1123
Tyr Ser Thr Ile Asn Lys Ile Pro Gln Leu Thr Leu Asn Phe Ser Glu
                330                 335                 340

gtc aaa act gag aaa aag aat tca tca cct cct tct tca gat aaa acc     1171
Val Lys Thr Glu Lys Lys Asn Ser Ser Pro Pro Ser Ser Asp Lys Thr
            345                 350                 355

att att gca ccc aag gtt aaa gat cga aca cac aat gtg act gag aaa     1219
Ile Ile Ala Pro Lys Val Lys Asp Arg Thr His Asn Val Thr Glu Lys
        360                 365                 370

gtg acc cag gtt ctc tct tta gga gca gat gtc cta cct gaa tac aaa     1267
Val Thr Gln Val Leu Ser Leu Gly Ala Asp Val Leu Pro Glu Tyr Lys
    375                 380                 385

ctg cag aca cca cgc atc aac aag ttt acg ata ttg cac tac agc cct     1315
Leu Gln Thr Pro Arg Ile Asn Lys Phe Thr Ile Leu His Tyr Ser Pro
390                 395                 400                 405

ttc aag gca gtc tgg gac tgg ctt atc ctg ctg ttg gtc ata tac act     1363
Phe Lys Ala Val Trp Asp Trp Leu Ile Leu Leu Leu Val Ile Tyr Thr
                410                 415                 420

gct ata ttt act ccc tac tct gca gcc ttc ctc ctc aat gac aga gaa     1411
Ala Ile Phe Thr Pro Tyr Ser Ala Ala Phe Leu Leu Asn Asp Arg Glu
```

-continued

```
            425                 430                 435

gaa cag aag aga cga gaa tgt ggc tat tct tgt agc cct ttg aat gtg     1459
Glu Gln Lys Arg Arg Glu Cys Gly Tyr Ser Cys Ser Pro Leu Asn Val
        440                 445                 450

gta gac ttg att gtg gat att atg ttt atc ata gat att tta ata aac     1507
Val Asp Leu Ile Val Asp Ile Met Phe Ile Ile Asp Ile Leu Ile Asn
    455                 460                 465

ttc aga aca aca tat gtt aat cag aat gaa gaa gtg gta agt gat ccc     1555
Phe Arg Thr Thr Tyr Val Asn Gln Asn Glu Glu Val Val Ser Asp Pro
470                 475                 480                 485

gcc aaa ata gca ata cac tac ttc aaa ggc tgg ttc ctg att gac atg     1603
Ala Lys Ile Ala Ile His Tyr Phe Lys Gly Trp Phe Leu Ile Asp Met
                490                 495                 500

gtt gca gca att cct ttt gac ttg ctg att ttt gga tca ggt tct gat     1651
Val Ala Ala Ile Pro Phe Asp Leu Leu Ile Phe Gly Ser Gly Ser Asp
            505                 510                 515

gag aca aca aca tta att ggt ctt ttg aag act gcc cga ctc ctc cgt     1699
Glu Thr Thr Thr Leu Ile Gly Leu Leu Lys Thr Ala Arg Leu Leu Arg
        520                 525                 530

ctt gtg cgc gtg gcc agg aaa ctg gat cga tat tca gaa tat ggc gct     1747
Leu Val Arg Val Ala Arg Lys Leu Asp Arg Tyr Ser Glu Tyr Gly Ala
    535                 540                 545

gct gtt cta atg ctc tca atg tgc atc ttt gcc ctg aat gca cac tgg     1795
Ala Val Leu Met Leu Ser Met Cys Ile Phe Ala Leu Asn Ala His Trp
550                 555                 560                 565

ctg gct tgc att tgg tat gcg att ggg aat gta gaa agg cct tac ctg     1843
Leu Ala Cys Ile Trp Tyr Ala Ile Gly Asn Val Glu Arg Pro Tyr Leu
                570                 575                 580

act gac aaa atc gga tgg ttg gat tcc tta gga cag caa att ggg aaa     1891
Thr Asp Lys Ile Gly Trp Leu Asp Ser Leu Gly Gln Gln Ile Gly Lys
            585                 590                 595

cgt tac aat gac agt gac tca agt tct gga cca tcc att aaa gac aaa     1939
Arg Tyr Asn Asp Ser Asp Ser Ser Ser Gly Pro Ser Ile Lys Asp Lys
        600                 605                 610

tac gtc aca gca ctt tat ttt acc ttc agc agt tta acc agt gta gga     1987
Tyr Val Thr Ala Leu Tyr Phe Thr Phe Ser Ser Leu Thr Ser Val Gly
    615                 620                 625

ttc ggg aat gtg tct cct aac acg aat tcg gag aaa atc ttt tca att     2035
Phe Gly Asn Val Ser Pro Asn Thr Asn Ser Glu Lys Ile Phe Ser Ile
630                 635                 640                 645

tgt gtc atg ttg att ggc tca cta atg tat gca agc att ttt ggg aat     2083
Cys Val Met Leu Ile Gly Ser Leu Met Tyr Ala Ser Ile Phe Gly Asn
                650                 655                 660

gta tct gca att atc caa aga cta tac tcg gga act gcc agg tac cac     2131
Val Ser Ala Ile Ile Gln Arg Leu Tyr Ser Gly Thr Ala Arg Tyr His
            665                 670                 675

atg cag atg ctg cga gta aaa gag ttc att cgc ttt cac caa atc ccc     2179
Met Gln Met Leu Arg Val Lys Glu Phe Ile Arg Phe His Gln Ile Pro
        680                 685                 690

aac cct ctg agg caa cgt ctt gaa gaa tat ttc cag cac gca tgg act     2227
Asn Pro Leu Arg Gln Arg Leu Glu Glu Tyr Phe Gln His Ala Trp Thr
    695                 700                 705

tac acc aat ggc att gac atg aac atg gtc cta aag ggt ttc cca gaa     2275
Tyr Thr Asn Gly Ile Asp Met Asn Met Val Leu Lys Gly Phe Pro Glu
710                 715                 720                 725

tgc tta caa gca gac att tgt cta cat ctc aac cag aca ttg ctg caa     2323
Cys Leu Gln Ala Asp Ile Cys Leu His Leu Asn Gln Thr Leu Leu Gln
```

-continued

```
                730                 735                 740

aac tgc aaa gcc ttt cgg ggg gca agt aaa ggt tgc ctt aga gct ttg     2371
Asn Cys Lys Ala Phe Arg Gly Ala Ser Lys Gly Cys Leu Arg Ala Leu
            745                 750                 755

gca atg aag ttc aaa acc acc cat gca ctc caa gga gac acc ctc gtt     2419
Ala Met Lys Phe Lys Thr Thr His Ala Leu Gln Gly Asp Thr Leu Val
        760                 765                 770

cac tgt ggg gat gtc ctc act gca ctt tat ttc tta tcc aga ggc tcc     2467
His Cys Gly Asp Val Leu Thr Ala Leu Tyr Phe Leu Ser Arg Gly Ser
    775                 780                 785

att gaa atc tca aag aat gac atg gtg gtg gct att ctg gga aaa aat     2515
Ile Glu Ile Ser Lys Asn Asp Met Val Val Ala Ile Leu Gly Lys Asn
790                 795                 800                 805

gat ata ttt gga gaa atg gtt cat ctt tat gcc aaa cct gga aag tct     2563
Asp Ile Phe Gly Glu Met Val His Leu Tyr Ala Lys Pro Gly Lys Ser
                810                 815                 820

aat gca gat gta aga gcc ctc aca tac tgt gac ttg cat aag att cag     2611
Asn Ala Asp Val Arg Ala Leu Thr Tyr Cys Asp Leu His Lys Ile Gln
            825                 830                 835

cga gaa gac ttg tta gag gtt ttg gat atg tat cct gag ttt tct gat     2659
Arg Glu Asp Leu Leu Glu Val Leu Asp Met Tyr Pro Glu Phe Ser Asp
        840                 845                 850

cac ttt cta aca aac cta gag ttg act ttc aac cta agg cat gag agc     2707
His Phe Leu Thr Asn Leu Glu Leu Thr Phe Asn Leu Arg His Glu Ser
    855                 860                 865

gca aag gct gat ctc cta cga tca caa tcc atg aat gat tca gaa gga     2755
Ala Lys Ala Asp Leu Leu Arg Ser Gln Ser Met Asn Asp Ser Glu Gly
870                 875                 880                 885

gac aac tgt aaa cta aga aga agg aaa ttg tca ttt gaa agt gaa gga     2803
Asp Asn Cys Lys Leu Arg Arg Arg Lys Leu Ser Phe Glu Ser Glu Gly
                890                 895                 900

gag aaa gaa aac agt aca aat gat cct gaa gac tct gca gat acc ata     2851
Glu Lys Glu Asn Ser Thr Asn Asp Pro Glu Asp Ser Ala Asp Thr Ile
            905                 910                 915

aga cat tat cag agt tcc aag aga cac ttt gaa gag aaa aaa agc aga     2899
Arg His Tyr Gln Ser Ser Lys Arg His Phe Glu Glu Lys Lys Ser Arg
        920                 925                 930

tcc tca tct ttc atc tcc tcc att gat gat gaa caa aag ccg ctc ttc     2947
Ser Ser Ser Phe Ile Ser Ser Ile Asp Asp Glu Gln Lys Pro Leu Phe
    935                 940                 945

tca gga ata gta gac tct tct cca gga ata ggg aaa gca tct ggg ctc     2995
Ser Gly Ile Val Asp Ser Ser Pro Gly Ile Gly Lys Ala Ser Gly Leu
950                 955                 960                 965

gat ttt gaa gaa aca gtg ccc acc tca gga aga atg cac ata gat aaa     3043
Asp Phe Glu Glu Thr Val Pro Thr Ser Gly Arg Met His Ile Asp Lys
                970                 975                 980

aga agt cac tct tgc aaa gat atc act gac atg cga agc tgg gaa cga     3091
Arg Ser His Ser Cys Lys Asp Ile Thr Asp Met Arg Ser Trp Glu Arg
            985                 990                 995

gaa aat gca cat ccc cag cct gaa gac tcc agt cca tct gca ctt cag     3139
Glu Asn Ala His Pro Gln Pro Glu Asp Ser Ser Pro Ser Ala Leu Gln
        1000                1005                1010

cga gct gcc tgg ggt atc tct gaa acc gaa agc gac ctc acc tac ggg     3187
Arg Ala Ala Trp Gly Ile Ser Glu Thr Glu Ser Asp Leu Thr Tyr Gly
    1015                1020                1025

gaa gtg gaa caa aga tta gat ctg ctc cag gag caa ctt aac agg ctt     3235
Glu Val Glu Gln Arg Leu Asp Leu Leu Gln Glu Gln Leu Asn Arg Leu
```

-continued

```
1030                1035                1040                1045

gaa tcc caa atg acc act gac atc cag acc atc tta cag ttg ctg cag     3283
Glu Ser Gln Met Thr Thr Asp Ile Gln Thr Ile Leu Gln Leu Leu Gln
                1050                1055                1060

aaa caa acc act gtg gtc ccc cca gcc tac agt atg gta aca gca gga     3331
Lys Gln Thr Thr Val Val Pro Pro Ala Tyr Ser Met Val Thr Ala Gly
            1065                1070                1075

tca gaa tat cag aga ccc atc atc cag ctg atg aga acc agt caa ccg     3379
Ser Glu Tyr Gln Arg Pro Ile Ile Gln Leu Met Arg Thr Ser Gln Pro
        1080                1085                1090

gaa gca tcc atc aaa act gac cga agt ttc agc cct tcc tca caa tgt     3427
Glu Ala Ser Ile Lys Thr Asp Arg Ser Phe Ser Pro Ser Ser Gln Cys
    1095                1100                1105

cct gaa ttt cta gac ctt gaa aaa tct aaa ctt aaa tcc aaa gaa tcc     3475
Pro Glu Phe Leu Asp Leu Glu Lys Ser Lys Leu Lys Ser Lys Glu Ser
1110                1115                1120                1125

ctt tca agt ggg gtg cat ctg aac aca gct tca gaa gac aac ttg act     3523
Leu Ser Ser Gly Val His Leu Asn Thr Ala Ser Glu Asp Asn Leu Thr
                1130                1135                1140

tca ctt tta aaa caa gac agt gat ctc tct tta gag ctt cac ctg cgg     3571
Ser Leu Leu Lys Gln Asp Ser Asp Leu Ser Leu Glu Leu His Leu Arg
            1145                1150                1155

caa aga aaa act tac gtt cat cca att agg cat cct tct ttg cca gat     3619
Gln Arg Lys Thr Tyr Val His Pro Ile Arg His Pro Ser Leu Pro Asp
        1160                1165                1170

tca tcc cta agc act gta gga atc gtg ggt ctt cat agg cat gtt tct     3667
Ser Ser Leu Ser Thr Val Gly Ile Val Gly Leu His Arg His Val Ser
    1175                1180                1185

gat cct ggt ctt cca ggg aaa taa tcattttgta ctatttactc cacatacaat    3721
Asp Pro Gly Leu Pro Gly Lys  *
1190                1195

gtaagtgctt ttaatggctg ttttcctttt tctatttaaa tcctctctac ttgactcagg   3781

ggctcacaag gtaccattat atgcaaaagt actgtatatt ttcctaaatt gaagcttgta   3841

aggtaaaact gagcagttag gatgtaaata tacataagaa cttttggttc caaatgttaa   3901

aactgccagc atctcacggc accttatttt ttatttttat tttttaaatc acatgcatgt   3961

taggaaactc caatttctct tgcatggaga ctcctattta ctgcttttac taaaccagta   4021

cttcgttatg aaaatgcctt ccacgcaaat aagaaaccaa gggataaaac tgttcatgga   4081

tgcaactcaa attcagatga tcatcaaggc at                                 4113


<210> SEQ ID NO 82
<211> LENGTH: 1196
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 82

Met Pro Val Arg Arg Gly His Val Ala Pro Gln Asn Thr Phe Leu Gly
 1               5                  10                  15

Thr Ile Ile Arg Lys Phe Glu Gly Gln Asn Lys Lys Phe Ile Ile Ala
            20                  25                  30

Asn Ala Arg Val Gln Asn Cys Ala Ile Ile Tyr Cys Asn Asp Gly Phe
        35                  40                  45

Cys Glu Met Thr Gly Phe Ser Arg Pro Asp Val Met Gln Lys Pro Cys
    50                  55                  60
```

-continued

```
Thr Cys Asp Phe Leu His Gly Pro Glu Thr Lys Arg His Asp Ile Ala
65                  70                  75                  80

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

Thr Tyr Tyr His Lys Asn Gly Ser Thr Phe Ile Cys Asn Thr His Ile
            100                 105                 110

Ile Pro Val Lys Asn Gln Glu Gly Val Ala Met Met Phe Ile Ile Asn
        115                 120                 125

Phe Glu Tyr Val Thr Asp Asn Glu Asn Ala Ala Thr Pro Glu Arg Val
    130                 135                 140

Asn Pro Ile Leu Pro Ile Lys Thr Val Asn Arg Lys Phe Phe Gly Phe
145                 150                 155                 160

Lys Phe Pro Gly Leu Arg Val Leu Thr Tyr Arg Lys Gln Ser Leu Pro
                165                 170                 175

Gln Glu Asp Pro Asp Val Val Val Ile Asp Ser Ser Lys His Ser Asp
            180                 185                 190

Asp Ser Val Ala Met Lys His Phe Lys Ser Pro Thr Lys Glu Ser Cys
        195                 200                 205

Ser Pro Ser Glu Ala Asp Asp Thr Lys Ala Leu Ile Gln Pro Ser Lys
    210                 215                 220

Cys Ser Pro Leu Val Asn Ile Ser Gly Pro Leu Asp His Ser Ser Pro
225                 230                 235                 240

Lys Arg Gln Trp Asp Arg Leu Tyr Pro Asp Met Leu Gln Ser Ser Ser
                245                 250                 255

Gln Leu Ser His Ser Arg Ser Arg Glu Ser Leu Cys Ser Ile Arg Arg
            260                 265                 270

Ala Ser Ser Val His Asp Ile Glu Gly Phe Gly Val His Pro Lys Asn
        275                 280                 285

Ile Phe Arg Asp Arg His Ala Ser Glu Asp Asn Gly Arg Asn Val Lys
    290                 295                 300

Gly Pro Phe Asn His Ile Lys Ser Ser Leu Leu Gly Ser Thr Ser Asp
305                 310                 315                 320

Ser Asn Leu Asn Lys Tyr Ser Thr Ile Asn Lys Ile Pro Gln Leu Thr
                325                 330                 335

Leu Asn Phe Ser Glu Val Lys Thr Glu Lys Lys Asn Ser Ser Pro Pro
            340                 345                 350

Ser Ser Asp Lys Thr Ile Ile Ala Pro Lys Val Lys Asp Arg Thr His
        355                 360                 365

Asn Val Thr Glu Lys Val Thr Gln Val Leu Ser Leu Gly Ala Asp Val
    370                 375                 380

Leu Pro Glu Tyr Lys Leu Gln Thr Pro Arg Ile Asn Lys Phe Thr Ile
385                 390                 395                 400

Leu His Tyr Ser Pro Phe Lys Ala Val Trp Asp Trp Leu Ile Leu Leu
                405                 410                 415

Leu Val Ile Tyr Thr Ala Ile Phe Thr Pro Tyr Ser Ala Ala Phe Leu
            420                 425                 430

Leu Asn Asp Arg Glu Glu Gln Lys Arg Arg Glu Cys Gly Tyr Ser Cys
        435                 440                 445

Ser Pro Leu Asn Val Val Asp Leu Ile Val Asp Ile Met Phe Ile Ile
    450                 455                 460

Asp Ile Leu Ile Asn Phe Arg Thr Thr Tyr Val Asn Gln Asn Glu Glu
```

-continued

```
465                 470                 475                 480

Val Val Ser Asp Pro Ala Lys Ile Ala Ile His Tyr Phe Lys Gly Trp
                485                 490                 495

Phe Leu Ile Asp Met Val Ala Ala Ile Pro Phe Asp Leu Leu Ile Phe
            500                 505                 510

Gly Ser Gly Ser Asp Glu Thr Thr Thr Leu Ile Gly Leu Leu Lys Thr
        515                 520                 525

Ala Arg Leu Leu Arg Leu Val Arg Val Ala Arg Lys Leu Asp Arg Tyr
    530                 535                 540

Ser Glu Tyr Gly Ala Ala Val Leu Met Leu Ser Met Cys Ile Phe Ala
545                 550                 555                 560

Leu Asn Ala His Trp Leu Ala Cys Ile Trp Tyr Ala Ile Gly Asn Val
                565                 570                 575

Glu Arg Pro Tyr Leu Thr Asp Lys Ile Gly Trp Leu Asp Ser Leu Gly
            580                 585                 590

Gln Gln Ile Gly Lys Arg Tyr Asn Asp Ser Asp Ser Ser Ser Gly Pro
        595                 600                 605

Ser Ile Lys Asp Lys Tyr Val Thr Ala Leu Tyr Phe Thr Phe Ser Ser
    610                 615                 620

Leu Thr Ser Val Gly Phe Gly Asn Val Ser Pro Asn Thr Asn Ser Glu
625                 630                 635                 640

Lys Ile Phe Ser Ile Cys Val Met Leu Ile Gly Ser Leu Met Tyr Ala
                645                 650                 655

Ser Ile Phe Gly Asn Val Ser Ala Ile Ile Gln Arg Leu Tyr Ser Gly
            660                 665                 670

Thr Ala Arg Tyr His Met Gln Met Leu Arg Val Lys Glu Phe Ile Arg
        675                 680                 685

Phe His Gln Ile Pro Asn Pro Leu Arg Gln Arg Leu Glu Glu Tyr Phe
    690                 695                 700

Gln His Ala Trp Thr Tyr Thr Asn Gly Ile Asp Met Asn Met Val Leu
705                 710                 715                 720

Lys Gly Phe Pro Glu Cys Leu Gln Ala Asp Ile Cys Leu His Leu Asn
                725                 730                 735

Gln Thr Leu Leu Gln Asn Cys Lys Ala Phe Arg Gly Ala Ser Lys Gly
            740                 745                 750

Cys Leu Arg Ala Leu Ala Met Lys Phe Lys Thr Thr His Ala Leu Gln
        755                 760                 765

Gly Asp Thr Leu Val His Cys Gly Asp Val Leu Thr Ala Leu Tyr Phe
    770                 775                 780

Leu Ser Arg Gly Ser Ile Glu Ile Ser Lys Asn Asp Met Val Val Ala
785                 790                 795                 800

Ile Leu Gly Lys Asn Asp Ile Phe Gly Glu Met Val His Leu Tyr Ala
                805                 810                 815

Lys Pro Gly Lys Ser Asn Ala Asp Val Arg Ala Leu Thr Tyr Cys Asp
            820                 825                 830

Leu His Lys Ile Gln Arg Glu Asp Leu Leu Glu Val Leu Asp Met Tyr
        835                 840                 845

Pro Glu Phe Ser Asp His Phe Leu Thr Asn Leu Glu Leu Thr Phe Asn
    850                 855                 860

Leu Arg His Glu Ser Ala Lys Ala Asp Leu Leu Arg Ser Gln Ser Met
865                 870                 875                 880
```

-continued

```
Asn Asp Ser Glu Gly Asp Asn Cys Lys Leu Arg Arg Arg Lys Leu Ser
                885                 890                 895

Phe Glu Ser Glu Gly Glu Lys Glu Asn Ser Thr Asn Asp Pro Glu Asp
            900                 905                 910

Ser Ala Asp Thr Ile Arg His Tyr Gln Ser Ser Lys Arg His Phe Glu
        915                 920                 925

Glu Lys Lys Ser Arg Ser Ser Ser Phe Ile Ser Ser Ile Asp Asp Glu
    930                 935                 940

Gln Lys Pro Leu Phe Ser Gly Ile Val Asp Ser Ser Pro Gly Ile Gly
945                 950                 955                 960

Lys Ala Ser Gly Leu Asp Phe Glu Glu Thr Val Pro Thr Ser Gly Arg
                965                 970                 975

Met His Ile Asp Lys Arg Ser His Ser Cys Lys Asp Ile Thr Asp Met
            980                 985                 990

Arg Ser Trp Glu Arg Glu Asn Ala His Pro Gln Pro Glu Asp Ser Ser
        995                 1000                1005

Pro Ser Ala Leu Gln Arg Ala Ala Trp Gly Ile Ser Glu Thr Glu Ser
    1010                1015                1020

Asp Leu Thr Tyr Gly Glu Val Glu Gln Arg Leu Asp Leu Leu Gln Glu
1025                1030                1035                1040

Gln Leu Asn Arg Leu Glu Ser Gln Met Thr Thr Asp Ile Gln Thr Ile
                1045                1050                1055

Leu Gln Leu Leu Gln Lys Gln Thr Thr Val Val Pro Pro Ala Tyr Ser
            1060                1065                1070

Met Val Thr Ala Gly Ser Glu Tyr Gln Arg Pro Ile Ile Gln Leu Met
        1075                1080                1085

Arg Thr Ser Gln Pro Glu Ala Ser Ile Lys Thr Asp Arg Ser Phe Ser
    1090                1095                1100

Pro Ser Ser Gln Cys Pro Glu Phe Leu Asp Leu Glu Lys Ser Lys Leu
1105                1110                1115                1120

Lys Ser Lys Glu Ser Leu Ser Ser Gly Val His Leu Asn Thr Ala Ser
                1125                1130                1135

Glu Asp Asn Leu Thr Ser Leu Leu Lys Gln Asp Ser Asp Leu Ser Leu
            1140                1145                1150

Glu Leu His Leu Arg Gln Arg Lys Thr Tyr Val His Pro Ile Arg His
        1155                1160                1165

Pro Ser Leu Pro Asp Ser Ser Leu Ser Thr Val Gly Ile Val Gly Leu
    1170                1175                1180

His Arg His Val Ser Asp Pro Gly Leu Pro Gly Lys
1185                1190                1195


<210> SEQ ID NO 83
<211> LENGTH: 3591
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(3591)

<400> SEQUENCE: 83

atg cct gtg cgc agg ggg cat gtg gca cca caa aat aca ttt ctg ggg       48
Met Pro Val Arg Arg Gly His Val Ala Pro Gln Asn Thr Phe Leu Gly
 1               5                  10                  15
```

-continued

```
acc atc att cgg aaa ttt gaa ggg caa aat aaa aaa ttt atc att gca       96
Thr Ile Ile Arg Lys Phe Glu Gly Gln Asn Lys Lys Phe Ile Ile Ala
            20                  25                  30

aat gcc aga gtg cag aac tgt gcc atc att tat tgc aac gat ggg ttc      144
Asn Ala Arg Val Gln Asn Cys Ala Ile Ile Tyr Cys Asn Asp Gly Phe
        35                  40                  45

tgt gag atg act ggt ttc tcc agg cca gat gtc atg caa aag cca tgc      192
Cys Glu Met Thr Gly Phe Ser Arg Pro Asp Val Met Gln Lys Pro Cys
    50                  55                  60

acc tgc gac ttt ctc cat gga ccc gag acc aag agg cat gat att gcc      240
Thr Cys Asp Phe Leu His Gly Pro Glu Thr Lys Arg His Asp Ile Ala
65                  70                  75                  80

caa att gcc cag gca ttg ctg ggg tca gaa gag agg aaa gtg gag gtc      288
Gln Ile Ala Gln Ala Leu Leu Gly Ser Glu Glu Arg Lys Val Glu Val
                85                  90                  95

acc tac tat cac aaa aat ggg tcc act ttt att tgt aac act cac ata      336
Thr Tyr Tyr His Lys Asn Gly Ser Thr Phe Ile Cys Asn Thr His Ile
            100                 105                 110

att cca gtg aaa aac caa gag ggc gtg gct atg atg ttc atc att aat      384
Ile Pro Val Lys Asn Gln Glu Gly Val Ala Met Met Phe Ile Ile Asn
        115                 120                 125

ttt gaa tat gtg acg gat aat gaa aac gct gcc acc cca gag agg gta      432
Phe Glu Tyr Val Thr Asp Asn Glu Asn Ala Ala Thr Pro Glu Arg Val
    130                 135                 140

aac cca ata tta cca atc aaa act gta aac cgg aaa ttt ttt ggg ttc      480
Asn Pro Ile Leu Pro Ile Lys Thr Val Asn Arg Lys Phe Phe Gly Phe
145                 150                 155                 160

aaa ttc cct ggt ctg aga gtt ctc act tac aga aag cag tcc tta cca      528
Lys Phe Pro Gly Leu Arg Val Leu Thr Tyr Arg Lys Gln Ser Leu Pro
                165                 170                 175

caa gaa gac ccc gat gtg gtg gtc atc gat tca tct aaa cac agt gat      576
Gln Glu Asp Pro Asp Val Val Val Ile Asp Ser Ser Lys His Ser Asp
            180                 185                 190

gat tca gta gcc atg aag cat ttt aag tct cct aca aaa gaa agc tgc      624
Asp Ser Val Ala Met Lys His Phe Lys Ser Pro Thr Lys Glu Ser Cys
        195                 200                 205

agc ccc tct gaa gca gat gac aca aaa gct ttg ata cag ccc agc aaa      672
Ser Pro Ser Glu Ala Asp Asp Thr Lys Ala Leu Ile Gln Pro Ser Lys
    210                 215                 220

tgt tct ccc ttg gtg aat ata tcc gga cct ctt gac cat tcc tct ccc      720
Cys Ser Pro Leu Val Asn Ile Ser Gly Pro Leu Asp His Ser Ser Pro
225                 230                 235                 240

aaa agg caa tgg gac cga ctc tac cct gac atg ctg cag tca agt tcc      768
Lys Arg Gln Trp Asp Arg Leu Tyr Pro Asp Met Leu Gln Ser Ser Ser
                245                 250                 255

cag ctg tcc cat tcc aga tca agg gaa agc tta tgt agt ata cgg aga      816
Gln Leu Ser His Ser Arg Ser Arg Glu Ser Leu Cys Ser Ile Arg Arg
            260                 265                 270

gca tct tcg gtc cat gat ata gaa gga ttc ggc gtc cac ccc aag aac      864
Ala Ser Ser Val His Asp Ile Glu Gly Phe Gly Val His Pro Lys Asn
        275                 280                 285

ata ttt aga gac cga cat gcc agc gaa gac aat ggt cgc aat gtc aaa      912
Ile Phe Arg Asp Arg His Ala Ser Glu Asp Asn Gly Arg Asn Val Lys
    290                 295                 300

ggg cct ttt aat cat atc aag tca agc ctc ctg gga tcc aca tca gat      960
Gly Pro Phe Asn His Ile Lys Ser Ser Leu Leu Gly Ser Thr Ser Asp
305                 310                 315                 320
```

-continued

```
tca aac ctc aac aaa tac agc acc att aac aag att cca cag ctc act     1008
Ser Asn Leu Asn Lys Tyr Ser Thr Ile Asn Lys Ile Pro Gln Leu Thr
                325                 330                 335

ctg aat ttt tca gag gtc aaa act gag aaa aag aat tca tca cct cct     1056
Leu Asn Phe Ser Glu Val Lys Thr Glu Lys Lys Asn Ser Ser Pro Pro
            340                 345                 350

tct tca gat aaa acc att att gca ccc aag gtt aaa gat cga aca cac     1104
Ser Ser Asp Lys Thr Ile Ile Ala Pro Lys Val Lys Asp Arg Thr His
        355                 360                 365

aat gtg act gag aaa gtg acc cag gtt ctc tct tta gga gca gat gtc     1152
Asn Val Thr Glu Lys Val Thr Gln Val Leu Ser Leu Gly Ala Asp Val
    370                 375                 380

cta cct gaa tac aaa ctg cag aca cca cgc atc aac aag ttt acg ata     1200
Leu Pro Glu Tyr Lys Leu Gln Thr Pro Arg Ile Asn Lys Phe Thr Ile
385                 390                 395                 400

ttg cac tac agc cct ttc aag gca gtc tgg gac tgg ctt atc ctg ctg     1248
Leu His Tyr Ser Pro Phe Lys Ala Val Trp Asp Trp Leu Ile Leu Leu
                405                 410                 415

ttg gtc ata tac act gct ata ttt act ccc tac tct gca gcc ttc ctc     1296
Leu Val Ile Tyr Thr Ala Ile Phe Thr Pro Tyr Ser Ala Ala Phe Leu
            420                 425                 430

ctc aat gac aga gaa gaa cag aag aga cga gaa tgt ggc tat tct tgt     1344
Leu Asn Asp Arg Glu Glu Gln Lys Arg Arg Glu Cys Gly Tyr Ser Cys
        435                 440                 445

agc cct ttg aat gtg gta gac ttg att gtg gat att atg ttt atc ata     1392
Ser Pro Leu Asn Val Val Asp Leu Ile Val Asp Ile Met Phe Ile Ile
    450                 455                 460

gat att tta ata aac ttc aga aca aca tat gtt aat cag aat gaa gaa     1440
Asp Ile Leu Ile Asn Phe Arg Thr Thr Tyr Val Asn Gln Asn Glu Glu
465                 470                 475                 480

gtg gta agt gat ccc gcc aaa ata gca ata cac tac ttc aaa ggc tgg     1488
Val Val Ser Asp Pro Ala Lys Ile Ala Ile His Tyr Phe Lys Gly Trp
                485                 490                 495

ttc ctg att gac atg gtt gca gca att cct ttt gac ttg ctg att ttt     1536
Phe Leu Ile Asp Met Val Ala Ala Ile Pro Phe Asp Leu Leu Ile Phe
            500                 505                 510

gga tca ggt tct gat gag aca aca aca tta att ggt ctt ttg aag act     1584
Gly Ser Gly Ser Asp Glu Thr Thr Thr Leu Ile Gly Leu Leu Lys Thr
        515                 520                 525

gcc cga ctc ctc cgt ctt gtg cgc gtg gcc agg aaa ctg gat cga tat     1632
Ala Arg Leu Leu Arg Leu Val Arg Val Ala Arg Lys Leu Asp Arg Tyr
    530                 535                 540

tca gaa tat ggc gct gct gtt cta atg ctc tca atg tgc atc ttt gcc     1680
Ser Glu Tyr Gly Ala Ala Val Leu Met Leu Ser Met Cys Ile Phe Ala
545                 550                 555                 560

ctg aat gca cac tgg ctg gct tgc att tgg tat gcg att ggg aat gta     1728
Leu Asn Ala His Trp Leu Ala Cys Ile Trp Tyr Ala Ile Gly Asn Val
                565                 570                 575

gaa agg cct tac ctg act gac aaa atc gga tgg ttg gat tcc tta gga     1776
Glu Arg Pro Tyr Leu Thr Asp Lys Ile Gly Trp Leu Asp Ser Leu Gly
            580                 585                 590

cag caa att ggg aaa cgt tac aat gac agt gac tca agt tct gga cca     1824
Gln Gln Ile Gly Lys Arg Tyr Asn Asp Ser Asp Ser Ser Ser Gly Pro
        595                 600                 605

tcc att aaa gac aaa tac gtc aca gca ctt tat ttt acc ttc agc agt     1872
Ser Ile Lys Asp Lys Tyr Val Thr Ala Leu Tyr Phe Thr Phe Ser Ser
    610                 615                 620
```

-continued

```
tta acc agt gta gga ttc ggg aat gtg tct cct aac acg aat tcg gag     1920
Leu Thr Ser Val Gly Phe Gly Asn Val Ser Pro Asn Thr Asn Ser Glu
625                 630                 635                 640

aaa atc ttt tca att tgt gtc atg ttg att ggc tca cta atg tat gca     1968
Lys Ile Phe Ser Ile Cys Val Met Leu Ile Gly Ser Leu Met Tyr Ala
                645                 650                 655

agc att ttt ggg aat gta tct gca att atc caa aga cta tac tcg gga     2016
Ser Ile Phe Gly Asn Val Ser Ala Ile Ile Gln Arg Leu Tyr Ser Gly
            660                 665                 670

act gcc agg tac cac atg cag atg ctg cga gta aaa gag ttc att cgc     2064
Thr Ala Arg Tyr His Met Gln Met Leu Arg Val Lys Glu Phe Ile Arg
        675                 680                 685

ttt cac caa atc ccc aac cct ctg agg caa cgt ctt gaa gaa tat ttc     2112
Phe His Gln Ile Pro Asn Pro Leu Arg Gln Arg Leu Glu Glu Tyr Phe
    690                 695                 700

cag cac gca tgg act tac acc aat ggc att gac atg aac atg gtc cta     2160
Gln His Ala Trp Thr Tyr Thr Asn Gly Ile Asp Met Asn Met Val Leu
705                 710                 715                 720

aag ggt ttc cca gaa tgc tta caa gca gac att tgt cta cat ctc aac     2208
Lys Gly Phe Pro Glu Cys Leu Gln Ala Asp Ile Cys Leu His Leu Asn
                725                 730                 735

cag aca ttg ctg caa aac tgc aaa gcc ttt cgg ggg gca agt aaa ggt     2256
Gln Thr Leu Leu Gln Asn Cys Lys Ala Phe Arg Gly Ala Ser Lys Gly
            740                 745                 750

tgc ctt aga gct ttg gca atg aag ttc aaa acc acc cat gca ctc caa     2304
Cys Leu Arg Ala Leu Ala Met Lys Phe Lys Thr Thr His Ala Leu Gln
        755                 760                 765

gga gac acc ctc gtt cac tgt ggg gat gtc ctc act gca ctt tat ttc     2352
Gly Asp Thr Leu Val His Cys Gly Asp Val Leu Thr Ala Leu Tyr Phe
    770                 775                 780

tta tcc aga ggc tcc att gaa atc tca aag aat gac atg gtg gtg gct     2400
Leu Ser Arg Gly Ser Ile Glu Ile Ser Lys Asn Asp Met Val Val Ala
785                 790                 795                 800

att ctg gga aaa aat gat ata ttt gga gaa atg gtt cat ctt tat gcc     2448
Ile Leu Gly Lys Asn Asp Ile Phe Gly Glu Met Val His Leu Tyr Ala
                805                 810                 815

aaa cct gga aag tct aat gca gat gta aga gcc ctc aca tac tgt gac     2496
Lys Pro Gly Lys Ser Asn Ala Asp Val Arg Ala Leu Thr Tyr Cys Asp
            820                 825                 830

ttg cat aag att cag cga gaa gac ttg tta gag gtt ttg gat atg tat     2544
Leu His Lys Ile Gln Arg Glu Asp Leu Leu Glu Val Leu Asp Met Tyr
        835                 840                 845

cct gag ttt tct gat cac ttt cta aca aac cta gag ttg act ttc aac     2592
Pro Glu Phe Ser Asp His Phe Leu Thr Asn Leu Glu Leu Thr Phe Asn
    850                 855                 860

cta agg cat gag agc gca aag gct gat ctc cta cga tca caa tcc atg     2640
Leu Arg His Glu Ser Ala Lys Ala Asp Leu Leu Arg Ser Gln Ser Met
865                 870                 875                 880

aat gat tca gaa gga gac aac tgt aaa cta aga aga agg aaa ttg tca     2688
Asn Asp Ser Glu Gly Asp Asn Cys Lys Leu Arg Arg Arg Lys Leu Ser
                885                 890                 895

ttt gaa agt gaa gga gag aaa gaa aac agt aca aat gat cct gaa gac     2736
Phe Glu Ser Glu Gly Glu Lys Glu Asn Ser Thr Asn Asp Pro Glu Asp
            900                 905                 910

tct gca gat acc ata aga cat tat cag agt tcc aag aga cac ttt gaa     2784
Ser Ala Asp Thr Ile Arg His Tyr Gln Ser Ser Lys Arg His Phe Glu
        915                 920                 925
```

-continued

```
gag aaa aaa agc aga tcc tca tct ttc atc tcc tcc att gat gat gaa     2832
Glu Lys Lys Ser Arg Ser Ser Ser Phe Ile Ser Ser Ile Asp Asp Glu
    930                 935                 940

caa aag ccg ctc ttc tca gga ata gta gac tct tct cca gga ata ggg     2880
Gln Lys Pro Leu Phe Ser Gly Ile Val Asp Ser Ser Pro Gly Ile Gly
945                 950                 955                 960

aaa gca tct ggg ctc gat ttt gaa gaa aca gtg ccc acc tca gga aga     2928
Lys Ala Ser Gly Leu Asp Phe Glu Glu Thr Val Pro Thr Ser Gly Arg
                965                 970                 975

atg cac ata gat aaa aga agt cac tct tgc aaa gat atc act gac atg     2976
Met His Ile Asp Lys Arg Ser His Ser Cys Lys Asp Ile Thr Asp Met
            980                 985                 990

cga agc tgg gaa cga gaa aat gca cat ccc cag cct gaa gac tcc agt     3024
Arg Ser Trp Glu Arg Glu Asn Ala His Pro Gln Pro Glu Asp Ser Ser
        995                 1000                1005

cca tct gca ctt cag cga gct gcc tgg ggt atc tct gaa acc gaa agc     3072
Pro Ser Ala Leu Gln Arg Ala Ala Trp Gly Ile Ser Glu Thr Glu Ser
    1010                1015                1020

gac ctc acc tac ggg gaa gtg gaa caa aga tta gat ctg ctc cag gag     3120
Asp Leu Thr Tyr Gly Glu Val Glu Gln Arg Leu Asp Leu Leu Gln Glu
1025                1030                1035                1040

caa ctt aac agg ctt gaa tcc caa atg acc act gac atc cag acc atc     3168
Gln Leu Asn Arg Leu Glu Ser Gln Met Thr Thr Asp Ile Gln Thr Ile
                1045                1050                1055

tta cag ttg ctg cag aaa caa acc act gtg gtc ccc cca gcc tac agt     3216
Leu Gln Leu Leu Gln Lys Gln Thr Thr Val Val Pro Pro Ala Tyr Ser
            1060                1065                1070

atg gta aca gca gga tca gaa tat cag aga ccc atc atc cag ctg atg     3264
Met Val Thr Ala Gly Ser Glu Tyr Gln Arg Pro Ile Ile Gln Leu Met
        1075                1080                1085

aga acc agt caa ccg gaa gca tcc atc aaa act gac cga agt ttc agc     3312
Arg Thr Ser Gln Pro Glu Ala Ser Ile Lys Thr Asp Arg Ser Phe Ser
    1090                1095                1100

cct tcc tca caa tgt cct gaa ttt cta gac ctt gaa aaa tct aaa ctt     3360
Pro Ser Ser Gln Cys Pro Glu Phe Leu Asp Leu Glu Lys Ser Lys Leu
1105                1110                1115                1120

aaa tcc aaa gaa tcc ctt tca agt ggg gtg cat ctg aac aca gct tca     3408
Lys Ser Lys Glu Ser Leu Ser Ser Gly Val His Leu Asn Thr Ala Ser
                1125                1130                1135

gaa gac aac ttg act tca ctt tta aaa caa gac agt gat ctc tct tta     3456
Glu Asp Asn Leu Thr Ser Leu Leu Lys Gln Asp Ser Asp Leu Ser Leu
            1140                1145                1150

gag ctt cac ctg cgg caa aga aaa act tac gtt cat cca att agg cat     3504
Glu Leu His Leu Arg Gln Arg Lys Thr Tyr Val His Pro Ile Arg His
        1155                1160                1165

cct tct ttg cca gat tca tcc cta agc act gta gga atc gtg ggt ctt     3552
Pro Ser Leu Pro Asp Ser Ser Leu Ser Thr Val Gly Ile Val Gly Leu
    1170                1175                1180

cat agg cat gtt tct gat cct ggt ctt cca ggg aaa taa                 3591
His Arg His Val Ser Asp Pro Gly Leu Pro Gly Lys  *
1185                1190                1195
```

<210> SEQ ID NO 84
<211> LENGTH: 215
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

-continued

<400> SEQUENCE: 84

Ile Leu Phe Ile Leu Asp Leu Leu Phe Val Leu Leu Phe Leu Leu Glu
1 5 10 15

Ile Val Leu Lys Phe Ile Ala Tyr Gly Leu Lys Ser Thr Ser Asn Ala
20 25 30

Lys Tyr Leu Lys Ser Ile Phe Asn Ile Leu Asp Leu Leu Ala Ile Leu
35 40 45

Pro Leu Leu Leu Leu Leu Val Leu Phe Leu Ser Gly Thr Glu Gln Val
50 55 60

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

Tyr Tyr Arg Ile Leu Arg Phe Leu Arg Leu Leu Arg Leu Leu Arg Leu
85 90 95

Leu Arg Leu Leu Arg Leu Leu Arg Arg Leu Glu Thr Leu Phe Glu Phe
100 105 110

Glu Leu Gly Thr Leu Ala Trp Ser Leu Gln Ser Leu Gly Leu Lys Ser
115 120 125

Ile Leu Arg Phe Leu Leu Leu Leu Leu Leu Leu Leu Ile Gly Phe Ser
130 135 140

Val Ile Gly Tyr Leu Leu Phe Lys Gly Tyr Glu Asp Leu Ser Glu Val
145 150 155 160

Asp Gly Asn Ser Glu Phe Ser Ser Tyr Phe Asp Ala Phe Tyr Phe Leu
165 170 175

Phe Val Thr Leu Thr Thr Val Gly Phe Gly Asp Leu Val Pro Val Gly
180 185 190

Ile Ile Phe Phe Val Leu Phe Phe Ile Ile Val Gly Leu Leu Leu Leu
195 200 205

Asn Leu Leu Ile Ala Val Ile
210 215

<210> SEQ ID NO 85
<211> LENGTH: 20
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 85

Ile Leu Tyr Trp Asn Ala Ala Ala Glu Glu Leu Thr Gly Leu Ser Arg
1 5 10 15

Glu Glu Val Ile
20

<210> SEQ ID NO 86
<211> LENGTH: 28
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 86

Thr Leu Glu Tyr Arg Asn Leu Arg Lys Asp Gly Ser Leu Ile Trp Val
1 5 10 15

Leu Val Ser Ala Ser Pro Ile Arg Asp Glu Asp Gly
20 25

-continued

<210> SEQ ID NO 87
<211> LENGTH: 120
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amino Acid Consensus Sequence

<400> SEQUENCE: 87

Ala Leu Glu Glu Arg Ser Tyr Pro Ala Gly Glu Val Ile Ile Arg Gln
1 5 10 15

Gly Asp Pro Gly Asp Ser Phe Tyr Ile Val Leu Ser Gly Glu Val Glu
20 25 30

Val Tyr Lys Leu Thr Glu Asp Gly Ala Arg Thr Pro Glu Val Ser Gln
35 40 45

Lys Gln Asp Thr Arg Glu Gln Val Val Ala Thr Leu Gly Pro Gly Asp
50 55 60

Phe Phe Gly Glu Leu Ala Leu Leu Thr Asn Asp Gly Asn Lys Asn Ala
65 70 75 80

Val Leu Pro Ser Leu Asp Gln Gly Ala Pro Arg Thr Ala Thr Val Arg
85 90 95

Ala Leu Thr Asp Ser Glu Leu Leu Arg Leu Asp Arg Glu Asp Phe Arg
100 105 110

Arg Leu Leu Gln Lys Tyr Pro Glu
115 120

<210> SEQ ID NO 88
<211> LENGTH: 31
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Coiled-coil segment

<400> SEQUENCE: 88

Ala Leu Glu Leu Ser Asn Glu Leu Ala Val Leu Ala Asn Ile Glu Lys
1 5 10 15

Glu Phe Lys Asn Asp Tyr Lys Lys Leu Ser Met Gln Cys Lys Asp
20 25 30

<210> SEQ ID NO 89
<211> LENGTH: 46
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Coiled-coil segment

<400> SEQUENCE: 89

Glu Val Asn Glu Gly Glu Leu Lys Glu Ile Lys Gln Asp Ile Ser Ser
1 5 10 15

Leu Arg Tyr Glu Leu Leu Glu Glu Lys Ser Gln Asn Thr Glu Asp Leu
20 25 30

Ala Glu Leu Ile Arg Glu Leu Gly Glu Lys Leu Ser Met Glu
35 40 45

<210> SEQ ID NO 90
<211> LENGTH: 22
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Leucine zipper pattern

-continued

<400> SEQUENCE: 90

Leu Val Pro Ser Pro Lys Ser Leu Phe Tyr Leu Leu Leu Lys Leu Lys
1 5 10 15

Lys Trp Ile Ser Glu Leu
20

<210> SEQ ID NO 91
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 91

Ser Ile Glu Glu
1

<210> SEQ ID NO 92
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 92

Ser Gln Ser Glu
1

<210> SEQ ID NO 93
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 93

Ser Ser His Asp
1

<210> SEQ ID NO 94
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 94

Ser Ser Glu Asp
1

<210> SEQ ID NO 95
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 95

Thr Ala Leu Glu
1

<210> SEQ ID NO 96

-continued

<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 96

Thr Ser Thr Asp
1

<210> SEQ ID NO 97
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 97

Ser Trp Met Glu
1

<210> SEQ ID NO 98
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 98

Ser Ile Ile Asp
1

<210> SEQ ID NO 99
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 99

Thr Leu Lys Asp
1

<210> SEQ ID NO 100
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 100

Thr Leu Gly Asp
1

<210> SEQ ID NO 101
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 101

Thr Val Lys Asp
1

-continued

<210> SEQ ID NO 102
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 102

Thr Thr Val Glu
1

<210> SEQ ID NO 103
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 103

Ser Phe Gln Glu
1

<210> SEQ ID NO 104
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 104

Ser Tyr Phe Glu
1

<210> SEQ ID NO 105
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 105

Ser His Glu Asp
1

<210> SEQ ID NO 106
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Casein kinase II phosphorylation site

<400> SEQUENCE: 106

Ser Ser Glu Asp
1

<210> SEQ ID NO 107
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: cAMP/cGMP-dependent protein kinase
phosphorylation site

<400> SEQUENCE: 107

-continued

Arg Arg Gln Thr
1

<210> SEQ ID NO 108
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: cAMP/cGMP-dependent protein kinase
phosphorylation site

<400> SEQUENCE: 108

Lys Lys Leu Ser
1

<210> SEQ ID NO 109
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Tyrosine phosphorylation site

<400> SEQUENCE: 109

Arg Arg Asn Glu Ser Gln Asp Tyr
1 5

<210> SEQ ID NO 110
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Tyrosine phosphorylation site

<400> SEQUENCE: 110

Arg Phe Leu Asp Ala Ala Glu Tyr
1 5

<210> SEQ ID NO 111
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Tyrosine phosphorylation site

<400> SEQUENCE: 111

Lys Phe Ile Glu Asn Ile Gly Tyr
1 5

<210> SEQ ID NO 112
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amidation site

<400> SEQUENCE: 112

Lys Gly Arg Arg
1

<210> SEQ ID NO 113
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Amidation site

-continued

<400> SEQUENCE: 113

Glu Gly Lys Arg
1

<210> SEQ ID NO 114
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: N-glycosylation site

<400> SEQUENCE: 114

Asn Glu Ser Gln
1

<210> SEQ ID NO 115
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: N-glycosylation site

<400> SEQUENCE: 115

Asn Leu Ser Arg
1

<210> SEQ ID NO 116
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: N-glycosylation site

<400> SEQUENCE: 116

Asn Leu Ser Arg
1

<210> SEQ ID NO 117
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: N-glycosylation site

<400> SEQUENCE: 117

Asn Leu Ser Gly
1

<210> SEQ ID NO 118
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: N-glycosylation site

<400> SEQUENCE: 118

Asn Glu Thr Ser
1

<210> SEQ ID NO 119
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence

-continued

<220> FEATURE:
<223> OTHER INFORMATION: N-glycosylation site

<400> SEQUENCE: 119

Asn Asp Thr Leu
1

<210> SEQ ID NO 120
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: N-glycosylation site

<400> SEQUENCE: 120

Asn Glu Ser Phe
1

<210> SEQ ID NO 121
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: N-glycosylation site

<400> SEQUENCE: 121

Asn Val Thr Met
1

<210> SEQ ID NO 122
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: N-glycosylation site

<400> SEQUENCE: 122

Asn Ser Ser Phe
1

<210> SEQ ID NO 123
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: N-myristoylation site

<400> SEQUENCE: 123

Gly Ala Ala Gly Ala Ala
1 5

<210> SEQ ID NO 124
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: N-myristoylation site

<400> SEQUENCE: 124

Gly Thr Arg Ser Ser His
1 5

<210> SEQ ID NO 125
<211> LENGTH: 6

-continued

```
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: N-myristoylation site

<400> SEQUENCE: 125

Gly Met Ile Trp Ala Glu
 1               5


<210> SEQ ID NO 126
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: N-myristoylation site

<400> SEQUENCE: 126

Gly Ala Lys Gln Asn Glu
 1               5


<210> SEQ ID NO 127
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: N-myristoylation site

<400> SEQUENCE: 127

Gly Ile Leu Gly Ser His
 1               5
```

What is claimed is:

1. An isolated 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid molecule selected from the group consisting of:

a) a nucleic acid molecule comprising a nucleotide sequence which is at least 60% identical to the nucleotide sequence of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83;

b) a nucleic acid molecule comprising a fragment of at least 15 nucleotides of the nucleotide sequence of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83;

c) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82;

d) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82;

e) a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, or a complement thereof, under stringent conditions;

f) a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83; and

g) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82.

2. The isolated nucleic acid molecule of claim **1**, which is the nucleotide sequence SEQ ID NO:1, 5, 13, 18, 21, 24, 27, 32, 35, 38, 53, 56, 60, 69, 72 or 81.

3. A host cell which contains the nucleic acid molecule of claim **1**.

4. An isolated 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide selected from the group consisting of:

a) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 60% identical to a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, or a complement thereof;

b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23,

24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83, or a complement thereof under stringent conditions;

c) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82; and

d) the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82.

5. An antibody which selectively binds to a polypeptide of claim **4**.

6. The polypeptide of claim **4**, further comprising heterologous amino acid sequences.

7. A method for producing a polypeptide selected from the group consisting of:

a) a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82;

b) a polypeptide comprising a fragment of the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82;

c) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, 3, 5, 7, 13, 15, 18, 20, 21, 23, 24, 26, 27, 29, 32, 34, 35, 37, 38, 40, 53, 55, 56, 58, 60, 62, 69, 71, 72, 74, 81 or 83; and

d) the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82;

comprising culturing the host cell of claim **3** under conditions in which the nucleic acid molecule is expressed.

8. A method for detecting the presence of a nucleic acid molecule of claim **1** or a polypeptide encoded by the nucleic acid molecule in a sample, comprising:

a) contacting the sample with a compound which selectively hybridizes to the nucleic acid molecule of claim **1** or binds to the polypeptide encoded by the nucleic acid molecule; and

b) determining whether the compound hybridizes to the nucleic acid or binds to the polypeptide in the sample.

9. A kit comprising a compound which selectively hybridizes to a nucleic acid molecule of claim **1** or binds to a polypeptide encoded by the nucleic acid molecule and instructions for use.

10. A method for identifying a compound which binds to a polypeptide or modulates the activity of the polypeptide of claim **4** comprising the steps of:

a) contacting a polypeptide, or a cell expressing a polypeptide of claim **4** with a test compound; and

b) determining whether the polypeptide binds to the test compound or determining the effect of the test compound on the activity of the polypeptide.

11. A method for modulating the activity of a polypeptide of claim **4** comprising contacting the polypeptide or a cell expressing the polypeptide with a compound which binds to the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

12. A method for identifying a compound capable of treating a disorder characterized by aberrant 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity, comprising assaying the ability of the compound to modulate 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid expression or 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide activity, thereby identifying a compound capable of treating a disorder characterized by aberrant 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity.

13. A method of identifying a nucleic acid molecule associated with a disorder characterized by aberrant 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity, comprising:

a) contacting a sample from a subject with a disorder characterized by aberrant 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity, comprising nucleic acid molecules with a hybridization probe comprising at least 25 contiguous nucleotides of SEQ ID NO:1, 5, 13, 18, 21, 24, 27, 32, 35, 38, 53, 56, 60, 69, 72 or 81 defined in claim **2**; and

b) detecting the presence of a nucleic acid molecule in the sample that hybridizes to the probe, thereby identifying a nucleic acid molecule associated with a disorder characterized by aberrant 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity.

14. A method of identifying a polypeptide associated with a disorder characterized by aberrant 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity, comprising:

a) contacting a sample comprising polypeptides with a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide defined in claim **4**; and

b) detecting the presence of a polypeptide in the sample that binds to the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 binding partner, thereby identifying the polypeptide associated with a disorder characterized by aberrant 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity.

15. A method of identifying a subject having a disorder characterized by aberrant 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity, comprising:

a) contacting a sample obtained from the subject comprising nucleic acid molecules with a hybridization probe comprising at least 25 contiguous nucleotides of SEQ ID NO:1, 5, 13, 18, 21, 24, 27, 32, 35, 38, 53, 56, 60, 69, 72 or 81 defined in claim **2**; and

b) detecting the presence of a nucleic acid molecule in the sample that hybridizes to the probe, thereby identifying a subject having a disorder characterized by aberrant 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity.

16. A method for treating a subject having a disorder characterized by aberrant 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity, or a subject at risk of developing a disorder characterized by aberrant 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 activity, comprising administering to the subject a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 modulator of the nucleic acid molecule defined in claim **1** or the polypeptide encoded by the nucleic acid molecule or contacting a cell with a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 modulator.

17. The method defined in claim **16** wherein said disorder is a cellular proliferative and/or differentiative disorder, brain disorder, blood vessel disorder, platelet disorder, breast disorder, colon disorder, kidney disorder, lung disorder, ovarian disorder, prostate disorder, pancreatic disorder, skeletal muscle disorder, testicular disorder, hormonal disorder, disorder associated with bone metabolism, immune e.g., inflammatory disorder, cardiovascular disorder, endothelial cell disorder, liver disorder, viral disease, pain, metabolic disorder, anemia, angiogenesic disorder, neoplastic disorder, endocrine disorder, neurological disorder or heart disorder.

18. The method of claim **16**, wherein the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 modulator is

a) a small molecule;

b) peptide;

c) phosphopeptide;

d) anti-18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 antibody;

e) a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide comprising the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, or a fragment thereof;

f) a 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 polypeptide comprising an amino acid sequence which is at least 90 percent identical to the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, wherein the percent identity is calculated using the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4; or

g) an isolated naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a complement of a nucleic acid molecule consisting of SEQ ID NO:1, 5, 13, 18, 21, 24, 27, 32, 35, 38, 53, 56, 60, 69, 72 or 81 at 6×SSC at 45° C., followed by one or more washes in 0.2×SSC, 0.1% SDS at 65° C.

19. The method of claim **16**, wherein the 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 modulator is

a) an antisense 18607, 15603, 69318, 12303, 48000, 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921 or 33751 nucleic acid molecule;

b) is a ribozyme;

c) the nucleotide sequence of SEQ ID NO:1, 5, 13, 18, 21, 24, 27, 32, 35, 38, 53, 56, 60, 69, 72 or 81 or a fragment thereof;

d) a nucleic acid molecule encoding a polypeptide comprising an amino acid sequence which is at least 90 percent identical to the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, wherein the percent identity is calculated using the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4;

e) a nucleic acid molecule encoding a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 6, 14, 19, 22, 25, 28, 33, 36, 39, 54, 57, 61, 70, 73 or 82, wherein the nucleic acid molecule which hybridizes to a complement of a nucleic acid molecule consisting of SEQ ID NO:1, 5, 13, 18, 21, 24, 27, 32, 35, 38, 53, 56, 60, 69, 72 or 81 at 6×SSC at 45° C., followed by one or more washes in 0.2×SSC, 0.1% SDS at 65° C.; or

f) a gene therapy vector.

* * * * *