

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

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
12 December 2002 (12.12.2002)

PCT

(10) International Publication Number  
**WO 02/099108 A2**

(51) International Patent Classification<sup>7</sup>: **C12N 15/12**,  
5/10, C07K 14/705, 16/18, G01N 33/50, A61K 45/00

(21) International Application Number: PCT/US01/44274

(22) International Filing Date:  
28 November 2001 (28.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/253,520 28 November 2000 (28.11.2000) US

(71) Applicant (*for all designated States except US*): **AMGEN, INC.** [US/US]; One Amgen Center Drive, Thousand Oaks, CA 91320 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **SHUTTER, John** [US/US]; 13175 Silver Creek Street, Moorpark, CA 93021 (US). **ULIAS, Laarni** [US/US]; 531 Innwood Road, Simi Valley, CA 93065 (US).

(74) Agent: **ZUHN, Donald, L.**; McDonnell Boehnen Hulbert & Berghoff, 300 South Wacker Drive, Suite 3200, Chicago, IL 60606 (US).

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

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

**Published:**

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

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

(54) Title: ATP-BINDING CASSETTE TRANSPORTER-LIKE MOLECULES AND USES THEREOF

(57) Abstract: The present invention provides ATP-Binding Cassette Transporter-Like (ABCL) polypeptides and nucleic acid molecules encoding the same. The invention also provides selective binding agents, vectors, host cells, and methods for producing ABCL polypeptides. The invention further provides pharmaceutical compositions and methods for the diagnosis, treatment, amelioration, and/or prevention of diseases, disorders, and conditions associated with ABCL polypeptides.



WO 02/099108 A2

## ATP-BINDING CASSETTE TRANSPORTER-LIKE MOLECULES AND USES THEREOF

This application claims the benefit of priority from U.S. Provisional Patent  
5 Application No. 60/253,520, filed on November 28, 2000, the disclosure of which is  
explicitly incorporated by reference herein.

### Field of the Invention

The present invention relates to ATP-Binding Cassette Transporter-Like  
10 (ABCL) polypeptides and nucleic acid molecules encoding the same. The invention  
also relates to selective binding agents, vectors, host cells, and methods for producing  
ABCL polypeptides. The invention further relates to pharmaceutical compositions  
and methods for the diagnosis, treatment, amelioration, and/or prevention of diseases,  
disorders, and conditions associated with ABCL polypeptides.

15

### Background of the Invention

Technical advances in the identification, cloning, expression, and  
manipulation of nucleic acid molecules and the deciphering of the human genome  
have greatly accelerated the discovery of novel therapeutics. Rapid nucleic acid  
20 sequencing techniques can now generate sequence information at unprecedented rates  
and, coupled with computational analyses, allow the assembly of overlapping  
sequences into partial and entire genomes and the identification of polypeptide-  
encoding regions. A comparison of a predicted amino acid sequence against a  
database compilation of known amino acid sequences allows one to determine the  
25 extent of homology to previously identified sequences and/or structural landmarks.  
The cloning and expression of a polypeptide-encoding region of a nucleic acid  
molecule provides a polypeptide product for structural and functional analyses. The  
manipulation of nucleic acid molecules and encoded polypeptides may confer  
advantageous properties on a product for use as a therapeutic.

30 In spite of the significant technical advances in genome research over the past  
decade, the potential for the development of novel therapeutics based on the human  
genome is still largely unrealized. Many genes encoding potentially beneficial

polypeptide therapeutics or those encoding polypeptides, which may act as “targets” for therapeutic molecules, have still not been identified. Accordingly, it is an object of the invention to identify novel polypeptides, and nucleic acid molecules encoding the same, which have diagnostic or therapeutic benefit.

5

The ATP-Binding Cassette (ABC) Transporter superfamily constitutes a large and diverse group of proteins that selectively mediate the movement of molecules across biological membranes. Over 100 ABC transporters have been identified, with the majority of these transporters being prokaryotic proteins. Although most members of this superfamily have some specificity for a particular substrate or group of related substrates, the number and types of substrates transported by the different members of the superfamily varies widely. For example, ABC transporters having substrate-specificity for proteins, sugars, peptides, polysaccharides, amino acids, and inorganic ions have been identified. Furthermore, some ABC transporters function to import substrates while other ABC transporters function to export substrates.

ABC transporters share a common organization consisting of four “core” domains: two transmembrane domains and two ATP-binding domains. All four domains may be present within the same multi-domain protein, or these domains may be expressed as separate polypeptides. The transmembrane domains, each consisting of six membrane-spanning helices, form the pathway through which substrates cross the membrane. The ATP-binding domains, or nucleotide binding folds (NBF), are located at the cytosolic face of the membrane. The NBF domains further comprise Walker A and B motifs, which are conserved among ABC transporters, regardless of the substrate specificity or species of origin of a particular ABC transporter. The substrate specificity of a particular transporter is ultimately determined by the transmembrane domains and the energy required to transport substrates is provided by ATP hydrolysis at the ATP-binding sites.

Several ABC transporters have been implicated in the pathogenesis of disease. Cystic fibrosis is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene, which encodes a chloride channel that regulates fluid secretion from exocrine tissues. The phenomena of multidrug resistance in tumor cells has been associated with the *mdr1* gene, which encodes an energy-

dependent efflux pump capable of removing toxic drugs from tumor cells. Tangier disease has been linked to mutations in the ABC1 gene, which has been shown to encode a protein responsible for the transport of unesterified cholesterol and phospholipids from certain cells. See Higgins, 1992, *Cell Biol.* 8:67-113; Young and  
5 Fielding, 1999, *Nat. Gen.* 22:316-18; Holland and Blight, 1999, *J. Mol. Biol.* 293:381-99.

#### Summary of the Invention

The present invention relates to novel ABCL nucleic acid molecules and  
10 encoded polypeptides.

The invention provides for an isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

- (a) the nucleotide sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ ID NO: 7;
- 15 (b) the nucleotide sequence of the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110, or PTA-3111;
- (c) a nucleotide sequence encoding the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;
- (d) a nucleotide sequence that hybridizes under at least moderately  
20 stringent conditions to the complement of the nucleotide sequence of any of (a) - (c);  
and
- (e) a nucleotide sequence complementary to the nucleotide sequence of any of (a) - (c).

25 The invention also provides for an isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence encoding a polypeptide that is at least about 70 percent identical to the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, wherein the encoded polypeptide has an activity of the  
30 polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;
- (b) a nucleotide sequence encoding an allelic variant or splice variant of the nucleotide sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ

ID NO: 7, the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110, or PTA-3111, or the nucleotide sequence of (a);

(c) a region of the nucleotide sequence of any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ ID NO: 7, the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110, or PTA-3111, or the nucleotide sequence of (a) or (b) encoding a polypeptide fragment of at least about 25 amino acid residues, wherein the polypeptide fragment has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, or is antigenic;

(d) a region of the nucleotide sequence of any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ ID NO: 7, the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110, or PTA-3111, or the nucleotide sequence of any of (a) - (c) comprising a fragment of at least about 16 nucleotides;

(e) a nucleotide sequence that hybridizes under at least moderately stringent conditions to the complement of the nucleotide sequence of any of (a) - (d); and

(f) a nucleotide sequence complementary to the nucleotide sequence of any of (a) - (d).

The invention further provides for an isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one conservative amino acid substitution, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

(b) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one amino acid insertion, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8

(c) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one amino acid deletion, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

(d) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 that has a C- and/or N- terminal truncation, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

5 (e) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions, C-terminal truncation, and N-terminal truncation, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID  
10 NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

(f) a nucleotide sequence of any of (a) - (e) comprising a fragment of at least about 16 nucleotides;

(g) a nucleotide sequence that hybridizes under at least moderately stringent conditions to the complement of the nucleotide sequence of any of (a) - (f);  
15 and

(h) a nucleotide sequence complementary to the nucleotide sequence of any of (a) - (e).

The present invention provides for an isolated polypeptide comprising an  
20 amino acid sequence selected from the group consisting of:

(a) the amino acid as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8; and

(b) the amino acid sequence encoded by the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110, or PTA-3111.

25

The invention also provides for an isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence as set forth in either SEQ ID NO: 3 or SEQ ID NO: 6, optionally further comprising an amino-terminal methionine;

30 (b) an amino acid sequence for an ortholog of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

- (c) an amino acid sequence that is at least about 70 percent identical to the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;
- 5 (d) a fragment of the amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 comprising at least about 25 amino acid residues, wherein the fragment has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, or is antigenic; and
- 10 (e) an amino acid sequence for an allelic variant or splice variant of the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110, or PTA-3111, or the amino acid sequence of any of (a) - (c).

The invention further provides for an isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

15

- (a) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one conservative amino acid substitution, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;
- 20 (b) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one amino acid insertion, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;
- (c) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one amino acid deletion, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;
- 25 (d) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 that has a C- and/or N- terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8; and
- 30

(e) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions, C-terminal truncation, and N-terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8.

Also provided are fusion polypeptides comprising ABCL amino acid sequences.

10 The present invention also provides for an expression vector comprising the isolated nucleic acid molecules as set forth herein, recombinant host cells comprising the recombinant nucleic acid molecules as set forth herein, and a method of producing an ABCL polypeptide comprising culturing the host cells and optionally isolating the polypeptide so produced.

15 A transgenic non-human animal comprising a nucleic acid molecule encoding an ABCL polypeptide is also encompassed by the invention. The ABCL nucleic acid molecules are introduced into the animal in a manner that allows expression and increased levels of an ABCL polypeptide, which may include increased circulating levels. Alternatively, the ABCL nucleic acid molecules are introduced into the animal in a manner that prevents expression of endogenous ABCL polypeptide (*i.e.*,  
20 generates a transgenic animal possessing an ABCL polypeptide gene knockout). The transgenic non-human animal is preferably a mammal, and more preferably a rodent, such as a rat or a mouse.

Also provided are derivatives of the ABCL polypeptides of the present  
25 invention.

Additionally provided are selective binding agents such as antibodies and peptides capable of specifically binding the ABCL polypeptides of the invention. Such antibodies and peptides may be agonistic or antagonistic.

Pharmaceutical compositions comprising the nucleotides, polypeptides, or  
30 selective binding agents of the invention and one or more pharmaceutically acceptable formulation agents are also encompassed by the invention. The pharmaceutical compositions are used to provide therapeutically effective amounts of

the nucleotides or polypeptides of the present invention. The invention is also directed to methods of using the polypeptides, nucleic acid molecules, and selective binding agents.

5 The ABCL polypeptides and nucleic acid molecules of the present invention may be used to treat, prevent, ameliorate, and/or detect diseases and disorders, including those recited herein.

The present invention also provides a method of assaying test molecules to identify a test molecule that binds to an ABCL polypeptide. The method comprises contacting an ABCL polypeptide with a test molecule to determine the extent of  
10 binding of the test molecule to the polypeptide. The method further comprises determining whether such test molecules are agonists or antagonists of an ABCL polypeptide. The present invention further provides a method of testing the impact of molecules on the expression of ABCL polypeptide or on the activity of ABCL polypeptide.

15 Methods of regulating expression and modulating (*i.e.*, increasing or decreasing) levels of an ABCL polypeptide are also encompassed by the invention. One method comprises administering to an animal a nucleic acid molecule encoding an ABCL polypeptide. In another method, a nucleic acid molecule comprising elements that regulate or modulate the expression of an ABCL polypeptide may be  
20 administered. Examples of these methods include gene therapy, cell therapy, and anti-sense therapy as further described herein.

ABCL polypeptides can be used for identifying ligands thereof. Various forms of "expression cloning" have been used for cloning ligands for receptors (*See, e.g.*, Davis *et al.*, 1996, *Cell*, 87:1161-69). These and other ABCL ligand cloning  
25 experiments are described in greater detail herein. Isolation of the ABCL ligand(s) allows for the identification or development of novel agonists and/or antagonists of the ABCL signaling pathway. Such agonists and antagonists include ABCL ligand(s), anti-ABCL ligand antibodies and derivatives thereof, small molecules, or antisense oligonucleotides, any of which can be used for potentially treating one or  
30 more diseases or disorders, including those recited herein.

#### Brief Description of the Figures

Figures 1A-1K illustrate the nucleotide sequence of the murine ABCL gene (SEQ ID NO: 1) and the deduced amino acid sequence (SEQ ID NO: 2) of the polypeptide encoded by this gene. The predicted signal peptide is indicated (underlined);

- 5     Figures 2A-2K illustrate the nucleotide sequence of the human ABCL gene (SEQ ID NO: 4) and the deduced amino acid sequence (SEQ ID NO: 5) of the polypeptide encoded by this gene. The predicted signal peptide is indicated (underlined);

- 10    Figures 3A-3H illustrate the nucleotide sequence of the human ABCL1550 gene (SEQ ID NO: 7) and the deduced amino acid sequence (SEQ ID NO: 8) of the polypeptide encoded by this gene;

- 15    Figures 4A-4E illustrate the amino acid sequence alignment of murine ABC1 polypeptide (muABC1; GenBank Accession No. NP\_038482; SEQ ID NO: 9), murine ABCR polypeptide (muABCR; GenBank Accession No. NP\_031404; SEQ ID NO: 10), and murine ABCL polypeptide (muABCL; SEQ ID NO: 2);

- 20    Figures 5A-5E illustrate the amino acid sequence alignment of human ABC1 polypeptide (huABC1; GenBank Accession No. AAF86276; SEQ ID NO: 11), human ABCR polypeptide (huABCR; GenBank Accession No. CAA75729; SEQ ID NO: 12), and human ABCL polypeptide (huABCL; SEQ ID NO: 5).

#### Detailed Description of the Invention

- 25    The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described. All references cited in this application are expressly incorporated by reference herein.

#### Definitions

- 30    The terms "ABCL gene" or "ABCL nucleic acid molecule" or "ABCL polynucleotide" refer to a nucleic acid molecule comprising or consisting of a nucleotide sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ ID NO: 7, a nucleotide sequence encoding the polypeptide as set forth in any of SEQ ID

NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, a nucleotide sequence of the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110, or PTA-3111, and nucleic acid molecules as defined herein.

The term "ABCL polypeptide allelic variant" refers to one of several possible naturally occurring alternate forms of a gene occupying a given locus on a chromosome of an organism or a population of organisms.

The term "ABCL polypeptide splice variant" refers to a nucleic acid molecule, usually RNA, which is generated by alternative processing of intron sequences in an RNA transcript of ABCL polypeptide amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8.

The term "isolated nucleic acid molecule" refers to a nucleic acid molecule of the invention that (1) has been separated from at least about 50 percent of proteins, lipids, carbohydrates, or other materials with which it is naturally found when total nucleic acid is isolated from the source cells, (2) is not linked to all or a portion of a polynucleotide to which the "isolated nucleic acid molecule" is linked in nature, (3) is operably linked to a polynucleotide which it is not linked to in nature, or (4) does not occur in nature as part of a larger polynucleotide sequence. Preferably, the isolated nucleic acid molecule of the present invention is substantially free from any other contaminating nucleic acid molecule(s) or other contaminants that are found in its natural environment that would interfere with its use in polypeptide production or its therapeutic, diagnostic, prophylactic or research use.

The term "nucleic acid sequence" or "nucleic acid molecule" refers to a DNA or RNA sequence. The term encompasses molecules formed from any of the known base analogs of DNA and RNA such as, but not limited to 4-acetylcytosine, 8-hydroxy-N6-methyladenosine, aziridinyl-cytosine, pseudoisocytosine, 5-(carboxyhydroxymethyl) uracil, 5-fluorouracil, 5-bromouracil, 5-carboxymethylaminomethyl-2-thiouracil, 5-carboxy-methylaminomethyluracil, dihydrouracil, inosine, N6-iso-pentenyladenine, 1-methyladenine, 1-methylpseudouracil, 1-methylguanine, 1-methylinosine, 2,2-dimethyl-guanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-methyladenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxycarbonyl-methyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarbonyl-methyluracil, 5-

methoxyuracil, 2-methylthio-N<sup>6</sup>-isopentenyladenine, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, oxybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, N-uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, pseudouracil, queosine, 2-thiocytosine, and 2,6-diaminopurine.

The term “vector” is used to refer to any molecule (*e.g.*, nucleic acid, plasmid, or virus) used to transfer coding information to a host cell.

The term “expression vector” refers to a vector that is suitable for transformation of a host cell and contains nucleic acid sequences that direct and/or control the expression of inserted heterologous nucleic acid sequences. Expression includes, but is not limited to, processes such as transcription, translation, and RNA splicing, if introns are present.

The term “operably linked” is used herein to refer to an arrangement of flanking sequences wherein the flanking sequences so described are configured or assembled so as to perform their usual function. Thus, a flanking sequence operably linked to a coding sequence may be capable of effecting the replication, transcription and/or translation of the coding sequence. For example, a coding sequence is operably linked to a promoter when the promoter is capable of directing transcription of that coding sequence. A flanking sequence need not be contiguous with the coding sequence, so long as it functions correctly. Thus, for example, intervening untranslated yet transcribed sequences can be present between a promoter sequence and the coding sequence and the promoter sequence can still be considered “operably linked” to the coding sequence.

The term “host cell” is used to refer to a cell which has been transformed, or is capable of being transformed with a nucleic acid sequence and then of expressing a selected gene of interest. The term includes the progeny of the parent cell, whether or not the progeny is identical in morphology or in genetic make-up to the original parent, so long as the selected gene is present.

The term “ABCL polypeptide” refers to a polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 and related polypeptides. Related polypeptides include ABCL polypeptide fragments, ABCL polypeptide orthologs, ABCL polypeptide variants, and ABCL polypeptide

derivatives, which possess at least one activity of the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8. ABCL polypeptides may be mature polypeptides, as defined herein, and may or may not have an amino-terminal methionine residue, depending on the method by which they are prepared.

5           The term "ABCL polypeptide fragment" refers to a polypeptide that comprises a truncation at the amino-terminus (with or without a leader sequence) and/or a truncation at the carboxyl-terminus of the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8. The term "ABCL polypeptide fragment" also refers to amino-terminal and/or carboxyl-terminal truncations of  
10 ABCL polypeptide orthologs, ABCL polypeptide derivatives, or ABCL polypeptide variants, or to amino-terminal and/or carboxyl-terminal truncations of the polypeptides encoded by ABCL polypeptide allelic variants or ABCL polypeptide splice variants. ABCL polypeptide fragments may result from alternative RNA splicing or from *in vivo* protease activity. Membrane-bound forms of an ABCL  
15 polypeptide are also contemplated by the present invention. In preferred embodiments, truncations and/or deletions comprise about 10 amino acids, or about 20 amino acids, or about 50 amino acids, or about 75 amino acids, or about 100 amino acids, or more than about 100 amino acids. The polypeptide fragments so produced will comprise about 25 contiguous amino acids, or about 50 amino acids, or  
20 about 75 amino acids, or about 100 amino acids, or about 150 amino acids, or about 200 amino acids, or more than about 200 amino acids. Such ABCL polypeptide fragments may optionally comprise an amino-terminal methionine residue. It will be appreciated that such fragments can be used, for example, to generate antibodies to ABCL polypeptides.

25           The term "ABCL polypeptide ortholog" refers to a polypeptide from another species that corresponds to ABCL polypeptide amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8. For example, mouse and human ABCL polypeptides are considered orthologs of each other.

          The term "ABCL polypeptide variants" refers to ABCL polypeptides  
30 comprising amino acid sequences having one or more amino acid sequence substitutions, deletions (such as internal deletions and/or ABCL polypeptide fragments), and/or additions (such as internal additions and/or ABCL fusion

polypeptides) as compared to the ABCL polypeptide amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 (with or without a leader sequence). Variants may be naturally occurring (*e.g.*, ABCL polypeptide allelic variants, ABCL polypeptide orthologs, and ABCL polypeptide splice variants) or artificially constructed. Such ABCL polypeptide variants may be prepared from the corresponding nucleic acid molecules having a DNA sequence that varies accordingly from the DNA sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ ID NO: 7. In preferred embodiments, the variants have from 1 to 3, or from 1 to 5, or from 1 to 10, or from 1 to 15, or from 1 to 20, or from 1 to 25, or from 1 to 50, or from 1 to 75, or from 1 to 100, or more than 100 amino acid substitutions, insertions, additions and/or deletions, wherein the substitutions may be conservative, or non-conservative, or any combination thereof.

The term "ABCL polypeptide derivatives" refers to the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, ABCL polypeptide fragments, ABCL polypeptide orthologs, or ABCL polypeptide variants, as defined herein, that have been chemically modified. The term "ABCL polypeptide derivatives" also refers to the polypeptides encoded by ABCL polypeptide allelic variants or ABCL polypeptide splice variants, as defined herein, that have been chemically modified.

The term "mature ABCL polypeptide" refers to an ABCL polypeptide lacking a leader sequence. A mature ABCL polypeptide may also include other modifications such as proteolytic processing of the amino-terminus (with or without a leader sequence) and/or the carboxyl-terminus, cleavage of a smaller polypeptide from a larger precursor, N-linked and/or O-linked glycosylation, and the like. Exemplary mature ABCL polypeptides are depicted by the amino acid sequences of SEQ ID NO: 3 and SEQ ID NO: 6.

The term "ABCL fusion polypeptide" refers to a fusion of one or more amino acids (such as a heterologous protein or peptide) at the amino- or carboxyl-terminus of the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, ABCL polypeptide fragments, ABCL polypeptide orthologs, ABCL polypeptide variants, or ABCL derivatives, as defined herein. The term "ABCL fusion polypeptide" also refers to a fusion of one or more amino acids at the amino-

or carboxyl-terminus of the polypeptide encoded by ABCL polypeptide allelic variants or ABCL polypeptide splice variants, as defined herein.

The term “biologically active ABCL polypeptides” refers to ABCL polypeptides having at least one activity characteristic of the polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8. In addition, an ABCL polypeptide may be active as an immunogen; that is, the ABCL polypeptide contains at least one epitope to which antibodies may be raised.

The term “isolated polypeptide” refers to a polypeptide of the present invention that (1) has been separated from at least about 50 percent of polynucleotides, lipids, carbohydrates, or other materials with which it is naturally found when isolated from the source cell, (2) is not linked (by covalent or noncovalent interaction) to all or a portion of a polypeptide to which the “isolated polypeptide” is linked in nature, (3) is operably linked (by covalent or noncovalent interaction) to a polypeptide with which it is not linked in nature, or (4) does not occur in nature. Preferably, the isolated polypeptide is substantially free from any other contaminating polypeptides or other contaminants that are found in its natural environment that would interfere with its therapeutic, diagnostic, prophylactic or research use.

The term “identity,” as known in the art, refers to a relationship between the sequences of two or more polypeptide molecules or two or more nucleic acid molecules, as determined by comparing the sequences. In the art, “identity” also means the degree of sequence relatedness between nucleic acid molecules or polypeptides, as the case may be, as determined by the match between strings of two or more nucleotide or two or more amino acid sequences. “Identity” measures the percent of identical matches between the smaller of two or more sequences with gap alignments (if any) addressed by a particular mathematical model or computer program (*i.e.*, “algorithms”).

The term “similarity” is a related concept, but in contrast to “identity,” “similarity” refers to a measure of relatedness that includes both identical matches and conservative substitution matches. If two polypeptide sequences have, for example, 10/20 identical amino acids, and the remainder are all non-conservative substitutions, then the percent identity and similarity would both be 50%. If in the

same example, there are five more positions where there are conservative substitutions, then the percent identity remains 50%, but the percent similarity would be 75% (15/20). Therefore, in cases where there are conservative substitutions, the percent similarity between two polypeptides will be higher than the percent identity  
5 between those two polypeptides.

The term “naturally occurring” or “native” when used in connection with biological materials such as nucleic acid molecules, polypeptides, host cells, and the like, refers to materials which are found in nature and are not manipulated by man. Similarly, “non-naturally occurring” or “non-native” as used herein refers to a  
10 material that is not found in nature or that has been structurally modified or synthesized by man.

The terms “effective amount” and “therapeutically effective amount” each refer to the amount of an ABCL polypeptide or ABCL nucleic acid molecule used to support an observable level of one or more biological activities of the ABCL  
15 polypeptides as set forth herein.

The term “pharmaceutically acceptable carrier” or “physiologically acceptable carrier” as used herein refers to one or more formulation materials suitable for accomplishing or enhancing the delivery of the ABCL polypeptide, ABCL nucleic acid molecule, or ABCL selective binding agent as a pharmaceutical composition.

20 The term “antigen” refers to a molecule or a portion of a molecule capable of being bound by a selective binding agent, such as an antibody, and additionally capable of being used in an animal to produce antibodies capable of binding to an epitope of that antigen. An antigen may have one or more epitopes.

The term “selective binding agent” refers to a molecule or molecules having  
25 specificity for an ABCL polypeptide. As used herein, the terms, “specific” and “specificity” refer to the ability of the selective binding agents to bind to human ABCL polypeptides and not to bind to human non-ABCL polypeptides. It will be appreciated, however, that the selective binding agents may also bind orthologs of the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8,  
30 that is, interspecies versions thereof, such as mouse and rat ABCL polypeptides.

The term “transduction” is used to refer to the transfer of genes from one bacterium to another, usually by a phage. “Transduction” also refers to the acquisition and transfer of eukaryotic cellular sequences by retroviruses.

The term “transfection” is used to refer to the uptake of foreign or exogenous DNA by a cell, and a cell has been “transfected” when the exogenous DNA has been introduced inside the cell membrane. A number of transfection techniques are well known in the art and are disclosed herein. See, e.g., Graham *et al.*, 1973, *Virology* 52:456; Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratories, 1989); Davis *et al.*, *Basic Methods in Molecular Biology* (Elsevier, 1986); and Chu *et al.*, 1981, *Gene* 13:197. Such techniques can be used to introduce one or more exogenous DNA moieties into suitable host cells.

The term “transformation” as used herein refers to a change in a cell’s genetic characteristics, and a cell has been transformed when it has been modified to contain a new DNA. For example, a cell is transformed where it is genetically modified from its native state. Following transfection or transduction, the transforming DNA may recombine with that of the cell by physically integrating into a chromosome of the cell, may be maintained transiently as an episomal element without being replicated, or may replicate independently as a plasmid. A cell is considered to have been stably transformed when the DNA is replicated with the division of the cell.

20

#### Relatedness of Nucleic Acid Molecules and/or Polypeptides

It is understood that related nucleic acid molecules include allelic or splice variants of the nucleic acid molecule of any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ ID NO: 7, and include sequences which are complementary to any of the above nucleotide sequences. Related nucleic acid molecules also include a nucleotide sequence encoding a polypeptide comprising or consisting essentially of a substitution, modification, addition and/or deletion of one or more amino acid residues compared to the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8. Such related ABCL polypeptides may comprise, for example, an addition and/or a deletion of one or more N-linked or O-linked glycosylation sites or an addition and/or a deletion of one or more cysteine residues.

30

Related nucleic acid molecules also include fragments of ABCL nucleic acid molecules which encode a polypeptide of at least about 25 contiguous amino acids, or about 50 amino acids, or about 75 amino acids, or about 100 amino acids, or about 150 amino acids, or about 200 amino acids, or more than 200 amino acid residues of the ABCL polypeptide of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8.

In addition, related ABCL nucleic acid molecules also include those molecules which comprise nucleotide sequences which hybridize under moderately or highly stringent conditions as defined herein with the fully complementary sequence of the ABCL nucleic acid molecule of any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ ID NO: 7, or of a molecule encoding a polypeptide, which polypeptide comprises the amino acid sequence as shown in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, or of a nucleic acid fragment as defined herein, or of a nucleic acid fragment encoding a polypeptide as defined herein. Hybridization probes may be prepared using the ABCL sequences provided herein to screen cDNA, genomic or synthetic DNA libraries for related sequences. Regions of the DNA and/or amino acid sequence of ABCL polypeptide that exhibit significant identity to known sequences are readily determined using sequence alignment algorithms as described herein and those regions may be used to design probes for screening.

The term "highly stringent conditions" refers to those conditions that are designed to permit hybridization of DNA strands whose sequences are highly complementary, and to exclude hybridization of significantly mismatched DNAs. Hybridization stringency is principally determined by temperature, ionic strength, and the concentration of denaturing agents such as formamide. Examples of "highly stringent conditions" for hybridization and washing are 0.015 M sodium chloride, 0.0015 M sodium citrate at 65-68°C or 0.015 M sodium chloride, 0.0015 M sodium citrate, and 50% formamide at 42°C. See Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual* (2nd ed., Cold Spring Harbor Laboratory, 1989); Anderson *et al.*, *Nucleic Acid Hybridisation: A Practical Approach* Ch. 4 (IRL Press Limited).

More stringent conditions (such as higher temperature, lower ionic strength, higher formamide, or other denaturing agent) may also be used – however, the rate of hybridization will be affected. Other agents may be included in the hybridization and

washing buffers for the purpose of reducing non-specific and/or background hybridization. Examples are 0.1% bovine serum albumin, 0.1% polyvinylpyrrolidone, 0.1% sodium pyrophosphate, 0.1% sodium dodecylsulfate, NaDodSO<sub>4</sub>, (SDS), ficoll, Denhardt's solution, sonicated salmon sperm DNA (or another non-complementary DNA), and dextran sulfate, although other suitable agents can also be used. The concentration and types of these additives can be changed without substantially affecting the stringency of the hybridization conditions. Hybridization experiments are usually carried out at pH 6.8-7.4; however, at typical ionic strength conditions, the rate of hybridization is nearly independent of pH. See Anderson *et al.*,  
10 *Nucleic Acid Hybridisation: A Practical Approach* Ch. 4 (IRL Press Limited).

Factors affecting the stability of DNA duplex include base composition, length, and degree of base pair mismatch. Hybridization conditions can be adjusted by one skilled in the art in order to accommodate these variables and allow DNAs of different sequence relatedness to form hybrids. The melting temperature of a perfectly matched DNA duplex can be estimated by the following equation:

$$T_m(^{\circ}\text{C}) = 81.5 + 16.6(\log[\text{Na}^+]) + 0.41(\% \text{G+C}) - 600/N - 0.72(\% \text{formamide})$$

where N is the length of the duplex formed, [Na<sup>+</sup>] is the molar concentration of the sodium ion in the hybridization or washing solution, %G+C is the percentage of (guanine+cytosine) bases in the hybrid. For imperfectly matched hybrids, the melting  
20 temperature is reduced by approximately 1°C for each 1% mismatch.

The term "moderately stringent conditions" refers to conditions under which a DNA duplex with a greater degree of base pair mismatching than could occur under "highly stringent conditions" is able to form. Examples of typical "moderately stringent conditions" are 0.015 M sodium chloride, 0.0015 M sodium citrate at 50-65°C or 0.015 M sodium chloride, 0.0015 M sodium citrate, and 20% formamide at  
25 37-50°C. By way of example, "moderately stringent conditions" of 50°C in 0.015 M sodium ion will allow about a 21% mismatch.

It will be appreciated by those skilled in the art that there is no absolute distinction between "highly stringent conditions" and "moderately stringent conditions." For example, at 0.015 M sodium ion (no formamide), the melting  
30 temperature of perfectly matched long DNA is about 71°C. With a wash at 65°C (at the same ionic strength), this would allow for approximately a 6% mismatch. To

capture more distantly related sequences, one skilled in the art can simply lower the temperature or raise the ionic strength.

A good estimate of the melting temperature in 1M NaCl\* for oligonucleotide probes up to about 20nt is given by:

5         $T_m = 2^{\circ}\text{C per A-T base pair} + 4^{\circ}\text{C per G-C base pair}$

\*The sodium ion concentration in 6X salt sodium citrate (SSC) is 1M. *See Suggs et al., Developmental Biology Using Purified Genes* 683 (Brown and Fox, eds., 1981).

High stringency washing conditions for oligonucleotides are usually at a temperature of 0-5°C below the  $T_m$  of the oligonucleotide in 6X SSC, 0.1% SDS.

10        In another embodiment, related nucleic acid molecules comprise or consist of a nucleotide sequence that is at least about 70 percent identical to the nucleotide sequence as shown in any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ ID NO: 7. In preferred embodiments, the nucleotide sequences are about 75 percent, or about 80 percent, or about 85 percent, or about 90 percent, or about 95, 96, 97, 98, or 99  
15        percent identical to the nucleotide sequence as shown in any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ ID NO: 7. Related nucleic acid molecules encode polypeptides possessing at least one activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8.

20        Differences in the nucleic acid sequence may result in conservative and/or non-conservative modifications of the amino acid sequence relative to the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8.

Conservative modifications to the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 (and the corresponding modifications to the encoding nucleotides) will produce a polypeptide having functional and chemical  
25        characteristics similar to those of ABCL polypeptides. In contrast, substantial modifications in the functional and/or chemical characteristics of ABCL polypeptides may be accomplished by selecting substitutions in the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 that differ significantly in their  
30        effect on maintaining (a) the structure of the molecular backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain.

For example, a “conservative amino acid substitution” may involve a substitution of a native amino acid residue with a nonnative residue such that there is little or no effect on the polarity or charge of the amino acid residue at that position. Furthermore, any native residue in the polypeptide may also be substituted with alanine, as has been previously described for “alanine scanning mutagenesis.”

Conservative amino acid substitutions also encompass non-naturally occurring amino acid residues that are typically incorporated by chemical peptide synthesis rather than by synthesis in biological systems. These include peptidomimetics, and other reversed or inverted forms of amino acid moieties.

Naturally occurring residues may be divided into classes based on common side chain properties:

- 1) hydrophobic: norleucine, Met, Ala, Val, Leu, Ile;
- 2) neutral hydrophilic: Cys, Ser, Thr;
- 3) acidic: Asp, Glu;
- 4) basic: Asn, Gln, His, Lys, Arg;
- 5) residues that influence chain orientation: Gly, Pro; and
- 6) aromatic: Trp, Tyr, Phe.

For example, non-conservative substitutions may involve the exchange of a member of one of these classes for a member from another class. Such substituted residues may be introduced into regions of the human ABCL polypeptide that are homologous with non-human ABCL polypeptides, or into the non-homologous regions of the molecule.

In making such changes, the hydropathic index of amino acids may be considered. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics. The hydropathic indices are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte *et al.*, 1982, *J.*

*Mol. Biol.* 157:105-31). It is known that certain amino acids may be substituted for other amino acids having a similar hydropathic index or score and still retain a similar biological activity. In making changes based upon the hydropathic index, the substitution of amino acids whose hydropathic indices are within  $\pm 2$  is preferred, those that are within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity, particularly where the biologically functionally equivalent protein or peptide thereby created is intended for use in immunological embodiments, as in the present case. The greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity and antigenicity, *i.e.*, with a biological property of the protein.

The following hydrophilicity values have been assigned to these amino acid residues: arginine (+3.0); lysine (+3.0); aspartate ( $+3.0 \pm 1$ ); glutamate ( $+3.0 \pm 1$ ); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline ( $-0.5 \pm 1$ ); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); and tryptophan (-3.4). In making changes based upon similar hydrophilicity values, the substitution of amino acids whose hydrophilicity values are within  $\pm 2$  is preferred, those that are within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred. One may also identify epitopes from primary amino acid sequences on the basis of hydrophilicity. These regions are also referred to as "epitopic core regions."

Desired amino acid substitutions (whether conservative or non-conservative) can be determined by those skilled in the art at the time such substitutions are desired. For example, amino acid substitutions can be used to identify important residues of the ABCL polypeptide, or to increase or decrease the affinity of the ABCL polypeptides described herein. Exemplary amino acid substitutions are set forth in Table I.

Table I

Amino Acid Substitutions

| Original Residues | Exemplary Substitutions                    | Preferred Substitutions |
|-------------------|--------------------------------------------|-------------------------|
| Ala               | Val, Leu, Ile                              | Val                     |
| Arg               | Lys, Gln, Asn                              | Lys                     |
| Asn               | Gln                                        | Gln                     |
| Asp               | Glu                                        | Glu                     |
| Cys               | Ser, Ala                                   | Ser                     |
| Gln               | Asn                                        | Asn                     |
| Glu               | Asp                                        | Asp                     |
| Gly               | Pro, Ala                                   | Ala                     |
| His               | Asn, Gln, Lys, Arg                         | Arg                     |
| Ile               | Leu, Val, Met, Ala,<br>Phe, Norleucine     | Leu                     |
| Leu               | Norleucine, Ile,<br>Val, Met, Ala, Phe     | Ile                     |
| Lys               | Arg, 1,4 Diamino-butyric<br>Acid, Gln, Asn | Arg                     |
| Met               | Leu, Phe, Ile                              | Leu                     |
| Phe               | Leu, Val, Ile, Ala,<br>Tyr                 | Leu                     |
| Pro               | Ala                                        | Gly                     |
| Ser               | Thr, Ala, Cys                              | Thr                     |
| Thr               | Ser                                        | Ser                     |
| Trp               | Tyr, Phe                                   | Tyr                     |
| Tyr               | Trp, Phe, Thr, Ser                         | Phe                     |
| Val               | Ile, Met, Leu, Phe,<br>Ala, Norleucine     | Leu                     |

A skilled artisan will be able to determine suitable variants of the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 using well-known techniques. For identifying suitable areas of the molecule that may be

changed without destroying biological activity, one skilled in the art may target areas not believed to be important for activity. For example, when similar polypeptides with similar activities from the same species or from other species are known, one skilled in the art may compare the amino acid sequence of an ABCL polypeptide to  
5 such similar polypeptides. With such a comparison, one can identify residues and portions of the molecules that are conserved among similar polypeptides. It will be appreciated that changes in areas of the ABCL molecule that are not conserved relative to such similar polypeptides would be less likely to adversely affect the biological activity and/or structure of an ABCL polypeptide. One skilled in the art  
10 would also know that, even in relatively conserved regions, one may substitute chemically similar amino acids for the naturally occurring residues while retaining activity (conservative amino acid residue substitutions). Therefore, even areas that may be important for biological activity or for structure may be subject to conservative amino acid substitutions without destroying the biological activity or  
15 without adversely affecting the polypeptide structure.

Additionally, one skilled in the art can review structure-function studies identifying residues in similar polypeptides that are important for activity or structure. In view of such a comparison, one can predict the importance of amino acid residues in an ABCL polypeptide that correspond to amino acid residues that are important for  
20 activity or structure in similar polypeptides. One skilled in the art may opt for chemically similar amino acid substitutions for such predicted important amino acid residues of ABCL polypeptides.

One skilled in the art can also analyze the three-dimensional structure and amino acid sequence in relation to that structure in similar polypeptides. In view of  
25 such information, one skilled in the art may predict the alignment of amino acid residues of ABCL polypeptide with respect to its three dimensional structure. One skilled in the art may choose not to make radical changes to amino acid residues predicted to be on the surface of the protein, since such residues may be involved in important interactions with other molecules. Moreover, one skilled in the art may  
30 generate test variants containing a single amino acid substitution at each amino acid residue. The variants could be screened using activity assays known to those with skill in the art. Such variants could be used to gather information about suitable

variants. For example, if one discovered that a change to a particular amino acid residue resulted in destroyed, undesirably reduced, or unsuitable activity, variants with such a change would be avoided. In other words, based on information gathered from such routine experiments, one skilled in the art can readily determine the amino acids where further substitutions should be avoided either alone or in combination with other mutations.

A number of scientific publications have been devoted to the prediction of secondary structure. See Moulton, 1996, *Curr. Opin. Biotechnol.* 7:422-27; Chou *et al.*, 1974, *Biochemistry* 13:222-45; Chou *et al.*, 1974, *Biochemistry* 113:211-22; Chou *et al.*, 1978, *Adv. Enzymol. Relat. Areas Mol. Biol.* 47:45-48; Chou *et al.*, 1978, *Ann. Rev. Biochem.* 47:251-276; and Chou *et al.*, 1979, *Biophys. J.* 26:367-84. Moreover, computer programs are currently available to assist with predicting secondary structure. One method of predicting secondary structure is based upon homology modeling. For example, two polypeptides or proteins that have a sequence identity of greater than 30%, or similarity greater than 40%, often have similar structural topologies. The recent growth of the protein structural database (PDB) has provided enhanced predictability of secondary structure, including the potential number of folds within the structure of a polypeptide or protein. See Holm *et al.*, 1999, *Nucleic Acids Res.* 27:244-47. It has been suggested that there are a limited number of folds in a given polypeptide or protein and that once a critical number of structures have been resolved, structural prediction will become dramatically more accurate (Brenner *et al.*, 1997, *Curr. Opin. Struct. Biol.* 7:369-76).

Additional methods of predicting secondary structure include "threading" (Jones, 1997, *Curr. Opin. Struct. Biol.* 7:377-87; Sippl *et al.*, 1996, *Structure* 4:15-19), "profile analysis" (Bowie *et al.*, 1991, *Science*, 253:164-70; Gribskov *et al.*, 1990, *Methods Enzymol.* 183:146-59; Gribskov *et al.*, 1987, *Proc. Nat. Acad. Sci. U.S.A.* 84:4355-58), and "evolutionary linkage" (See Holm *et al.*, *supra*, and Brenner *et al.*, *supra*).

Preferred ABCL polypeptide variants include glycosylation variants wherein the number and/or type of glycosylation sites have been altered compared to the amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8. In one embodiment, ABCL polypeptide variants comprise a greater or a lesser

number of N-linked glycosylation sites than the amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8. An N-linked glycosylation site is characterized by the sequence: Asn-X-Ser or Asn-X-Thr, wherein the amino acid residue designated as X may be any amino acid residue except proline. The substitution of amino acid residues to create this sequence provides a potential new site for the addition of an N-linked carbohydrate chain. Alternatively, substitutions that eliminate this sequence will remove an existing N-linked carbohydrate chain. Also provided is a rearrangement of N-linked carbohydrate chains wherein one or more N-linked glycosylation sites (typically those that are naturally occurring) are eliminated and one or more new N-linked sites are created. Additional preferred ABCL variants include cysteine variants, wherein one or more cysteine residues are deleted or substituted with another amino acid (*e.g.*, serine) as compared to the amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8. Cysteine variants are useful when ABCL polypeptides must be refolded into a biologically active conformation such as after the isolation of insoluble inclusion bodies. Cysteine variants generally have fewer cysteine residues than the native protein, and typically have an even number to minimize interactions resulting from unpaired cysteines.

In other embodiments, ABCL polypeptide variants comprise an amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one amino acid insertion and wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, or an amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one amino acid deletion and wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8. ABCL polypeptide variants also comprise an amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 wherein the polypeptide has a carboxyl- and/or amino-terminal truncation and further wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8. ABCL polypeptide variants further comprise an amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one modification selected from the group consisting of amino acid

substitutions, amino acid insertions, amino acid deletions, carboxyl-terminal truncations, and amino-terminal truncations and wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8.

5           In further embodiments, ABCL polypeptide variants comprise an amino acid sequence that is at least about 70 percent identical to the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8. In preferred embodiments, ABCL polypeptide variants comprise an amino acid sequence that is at least about 75 percent, or about 80 percent, or about 85 percent, or about 90 percent,  
10       or about 95, 96, 97, 98, or 99 percent identical percent to the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8. ABCL polypeptide variants possess at least one activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8.

          In addition, the polypeptide comprising the amino acid sequence of any of  
15       SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, or other ABCL polypeptide, may be fused to a homologous polypeptide to form a homodimer or to a heterologous polypeptide to form a heterodimer. Heterologous peptides and polypeptides include, but are not limited to: an epitope to allow for the detection and/or isolation of an ABCL fusion polypeptide; a transmembrane receptor protein or a portion thereof,  
20       such as an extracellular domain or a transmembrane and intracellular domain; a ligand or a portion thereof which binds to a transmembrane receptor protein; an enzyme or portion thereof which is catalytically active; a polypeptide or peptide which promotes oligomerization, such as a leucine zipper domain; a polypeptide or peptide which increases stability, such as an immunoglobulin constant region; and a  
25       polypeptide which has a therapeutic activity different from the polypeptide comprising the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, or other ABCL polypeptide.

          Fusions can be made either at the amino-terminus or at the carboxyl-terminus of the polypeptide comprising the amino acid sequence set forth in any of SEQ ID  
30       NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, or other ABCL polypeptide. Fusions may be direct with no linker or adapter molecule or may be through a linker or adapter molecule. A linker or adapter molecule may be one or more amino acid residues,

typically from about 20 to about 50 amino acid residues. A linker or adapter molecule may also be designed with a cleavage site for a DNA restriction endonuclease or for a protease to allow for the separation of the fused moieties. It will be appreciated that once constructed, the fusion polypeptides can be derivatized according to the methods described herein.

In a further embodiment of the invention, the polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, or other ABCL polypeptide, is fused to one or more domains of an Fc region of human IgG. Antibodies comprise two functionally independent parts, a variable domain known as "Fab," that binds an antigen, and a constant domain known as "Fc," that is involved in effector functions such as complement activation and attack by phagocytic cells. An Fc has a long serum half-life, whereas an Fab is short-lived. Capon *et al.*, 1989, *Nature* 337:525-31. When constructed together with a therapeutic protein, an Fc domain can provide longer half-life or incorporate such functions as Fc receptor binding, protein A binding, complement fixation, and perhaps even placental transfer. *Id.* Table II summarizes the use of certain Fc fusions known in the art.

Table II  
Fc Fusion with Therapeutic Proteins

| Form of Fc                                         | Fusion partner       | Therapeutic implications                                | Reference                                                                                                                        |
|----------------------------------------------------|----------------------|---------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| IgG1                                               | N-terminus of CD30-L | Hodgkin's disease; anaplastic lymphoma; T-cell leukemia | U.S. Patent No. 5,480,981                                                                                                        |
| Murine Fcγ2a                                       | IL-10                | anti-inflammatory; transplant rejection                 | Zheng <i>et al.</i> , 1995, <i>J. Immunol.</i> 154:5590-600                                                                      |
| IgG1                                               | TNF receptor         | septic shock                                            | Fisher <i>et al.</i> , 1996, <i>N. Engl. J. Med.</i> 334:1697-1702; Van Zee <i>et al.</i> , 1996, <i>J. Immunol.</i> 156:2221-30 |
| IgG, IgA, IgM, or IgE (excluding the first domain) | TNF receptor         | inflammation, autoimmune disorders                      | U.S. Patent No. 5,808,029                                                                                                        |
| IgG1                                               | CD4 receptor         | AIDS                                                    | Capon <i>et al.</i> , 1989, <i>Nature</i> 337: 525-31                                                                            |
| IgG1,                                              | N-terminus           | anti-cancer, antiviral                                  | Harvill <i>et al.</i> , 1995,                                                                                                    |

|                       |                      |                                 |                                                 |
|-----------------------|----------------------|---------------------------------|-------------------------------------------------|
| IgG3                  | of IL-2              |                                 | <i>Immunotech.</i> 1:95-105                     |
| IgG1                  | C-terminus of OPG    | osteoarthritis;<br>bone density | WO 97/23614                                     |
| IgG1                  | N-terminus of leptin | anti-obesity                    | PCT/US 97/23183, filed<br>December 11, 1997     |
| Human Ig C $\gamma$ 1 | CTLA-4               | autoimmune disorders            | Linsley, 1991, <i>J. Exp. Med.</i> , 174:561-69 |

In one example, a human IgG hinge, CH2, and CH3 region may be fused at either the amino-terminus or carboxyl-terminus of the ABCL polypeptides using methods known to the skilled artisan. In another example, a human IgG hinge, CH2, and CH3 region may be fused at either the amino-terminus or carboxyl-terminus of an ABCL polypeptide fragment (e.g., the predicted extracellular portion of ABCL polypeptide).

The resulting ABCL fusion polypeptide may be purified by use of a Protein A affinity column. Peptides and proteins fused to an Fc region have been found to exhibit a substantially greater half-life *in vivo* than the unfused counterpart. Also, a fusion to an Fc region allows for dimerization/multimerization of the fusion polypeptide. The Fc region may be a naturally occurring Fc region, or may be altered to improve certain qualities, such as therapeutic qualities, circulation time, or reduced aggregation.

Identity and similarity of related nucleic acid molecules and polypeptides are readily calculated by known methods. Such methods include, but are not limited to those described in *Computational Molecular Biology* (A.M. Lesk, ed., Oxford University Press 1988); *Biocomputing: Informatics and Genome Projects* (D.W. Smith, ed., Academic Press 1993); *Computer Analysis of Sequence Data* (Part 1, A.M. Griffin and H.G. Griffin, eds., Humana Press 1994); G. von Heinle, *Sequence Analysis in Molecular Biology* (Academic Press 1987); *Sequence Analysis Primer* (M. Gribskov and J. Devereux, eds., M. Stockton Press 1991); and Carillo *et al.*, 1988, *SIAM J. Applied Math.*, 48:1073.

Preferred methods to determine identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are described in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences

include, but are not limited to, the GCG program package, including GAP (Devereux *et al.*, 1984, *Nucleic Acids Res.* 12:387; Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, and FASTA (Altschul *et al.*, 1990, *J. Mol. Biol.* 215:403-10). The BLASTX program is publicly available from the  
5 National Center for Biotechnology Information (NCBI) and other sources (Altschul *et al.*, *BLAST Manual* (NCB NLM NIH, Bethesda, MD); Altschul *et al.*, 1990, *supra*). The well-known Smith Waterman algorithm may also be used to determine identity.

Certain alignment schemes for aligning two amino acid sequences may result in the matching of only a short region of the two sequences, and this small aligned  
10 region may have very high sequence identity even though there is no significant relationship between the two full-length sequences. Accordingly, in a preferred embodiment, the selected alignment method (GAP program) will result in an alignment that spans at least 50 contiguous amino acids of the claimed polypeptide.

For example, using the computer algorithm GAP (Genetics Computer Group,  
15 University of Wisconsin, Madison, WI), two polypeptides for which the percent sequence identity is to be determined are aligned for optimal matching of their respective amino acids (the "matched span," as determined by the algorithm). A gap opening penalty (which is calculated as 3X the average diagonal; the "average diagonal" is the average of the diagonal of the comparison matrix being used; the  
20 "diagonal" is the score or number assigned to each perfect amino acid match by the particular comparison matrix) and a gap extension penalty (which is usually 0.1X the gap opening penalty), as well as a comparison matrix such as PAM 250 or BLOSUM 62 are used in conjunction with the algorithm. A standard comparison matrix is also used by the algorithm (*see Dayhoff et al.*, 5 *Atlas of Protein Sequence and Structure* (Supp. 3 1978)(PAM250 comparison matrix); Henikoff *et al.*, 1992, *Proc. Natl. Acad. Sci USA* 89:10915-19 (BLOSUM 62 comparison matrix)).  
25

Preferred parameters for polypeptide sequence comparison include the following:

30       Algorithm: Needleman and Wunsch, 1970, *J. Mol. Biol.* 48:443-53;  
      Comparison matrix: BLOSUM 62 (Henikoff *et al.*, *supra*);  
      Gap Penalty: 12

Gap Length Penalty: 4

Threshold of Similarity: 0

5 The GAP program is useful with the above parameters. The aforementioned parameters are the default parameters for polypeptide comparisons (along with no penalty for end gaps) using the GAP algorithm.

Preferred parameters for nucleic acid molecule sequence comparison include the following:

10       Algorithm: Needleman and Wunsch, *supra*;  
          Comparison matrix: matches = +10, mismatch = 0  
          Gap Penalty: 50  
          Gap Length Penalty: 3

15 The GAP program is also useful with the above parameters. The aforementioned parameters are the default parameters for nucleic acid molecule comparisons.

Other exemplary algorithms, gap opening penalties, gap extension penalties, comparison matrices, and thresholds of similarity may be used, including those set forth in the Program Manual, Wisconsin Package, Version 9, September, 1997. The  
20 particular choices to be made will be apparent to those of skill in the art and will depend on the specific comparison to be made, such as DNA-to-DNA, protein-to-protein, protein-to-DNA; and additionally, whether the comparison is between given pairs of sequences (in which case GAP or BestFit are generally preferred) or between one sequence and a large database of sequences (in which case FASTA or BLASTA  
25 are preferred).

#### Nucleic Acid Molecules

The nucleic acid molecules encoding a polypeptide comprising the amino acid sequence of an ABCL polypeptide can readily be obtained in a variety of ways  
30 including, without limitation, chemical synthesis, cDNA or genomic library screening, expression library screening, and/or PCR amplification of cDNA.

Recombinant DNA methods used herein are generally those set forth in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 1989) and/or *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds., Green Publishers Inc. and Wiley and Sons 1994). The invention provides  
5 for nucleic acid molecules as described herein and methods for obtaining such molecules.

Where a gene encoding the amino acid sequence of an ABCL polypeptide has been identified from one species, all or a portion of that gene may be used as a probe to identify orthologs or related genes from the same species. The probes or primers  
10 may be used to screen cDNA libraries from various tissue sources believed to express the ABCL polypeptide. In addition, part or all of a nucleic acid molecule having the sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ ID NO: 7 may be used to screen a genomic library to identify and isolate a gene encoding the amino acid sequence of an ABCL polypeptide. Typically, conditions of moderate or high  
15 stringency will be employed for screening to minimize the number of false positives obtained from the screening.

Nucleic acid molecules encoding the amino acid sequence of ABCL polypeptides may also be identified by expression cloning which employs the detection of positive clones based upon a property of the expressed protein.  
20 Typically, nucleic acid libraries are screened by the binding an antibody or other binding partner (*e.g.*, receptor or ligand) to cloned proteins that are expressed and displayed on a host cell surface. The antibody or binding partner is modified with a detectable label to identify those cells expressing the desired clone.

Recombinant expression techniques conducted in accordance with the descriptions set forth below may be followed to produce these polynucleotides and to  
25 express the encoded polypeptides. For example, by inserting a nucleic acid sequence that encodes the amino acid sequence of an ABCL polypeptide into an appropriate vector, one skilled in the art can readily produce large quantities of the desired nucleotide sequence. The sequences can then be used to generate detection probes or  
30 amplification primers. Alternatively, a polynucleotide encoding the amino acid sequence of an ABCL polypeptide can be inserted into an expression vector. By

introducing the expression vector into an appropriate host, the encoded ABCL polypeptide may be produced in large amounts.

Another method for obtaining a suitable nucleic acid sequence is the polymerase chain reaction (PCR). In this method, cDNA is prepared from poly(A)+RNA or total RNA using the enzyme reverse transcriptase. Two primers, typically complementary to two separate regions of cDNA encoding the amino acid sequence of an ABCL polypeptide, are then added to the cDNA along with a polymerase such as *Taq* polymerase, and the polymerase amplifies the cDNA region between the two primers.

Another means of preparing a nucleic acid molecule encoding the amino acid sequence of an ABCL polypeptide is chemical synthesis using methods well known to the skilled artisan such as those described by Engels *et al.*, 1989, *Angew. Chem. Intl. Ed.* 28:716-34. These methods include, *inter alia*, the phosphotriester, phosphoramidite, and H-phosphonate methods for nucleic acid synthesis. A preferred method for such chemical synthesis is polymer-supported synthesis using standard phosphoramidite chemistry. Typically, the DNA encoding the amino acid sequence of an ABCL polypeptide will be several hundred nucleotides in length. Nucleic acids larger than about 100 nucleotides can be synthesized as several fragments using these methods. The fragments can then be ligated together to form the full-length nucleotide sequence of an ABCL gene. Usually, the DNA fragment encoding the amino-terminus of the polypeptide will have an ATG, which encodes a methionine residue. This methionine may or may not be present on the mature form of the ABCL polypeptide, depending on whether the polypeptide produced in the host cell is designed to be secreted from that cell. Other methods known to the skilled artisan may be used as well.

In certain embodiments, nucleic acid variants contain codons which have been altered for optimal expression of an ABCL polypeptide in a given host cell. Particular codon alterations will depend upon the ABCL polypeptide and host cell selected for expression. Such "codon optimization" can be carried out by a variety of methods, for example, by selecting codons which are preferred for use in highly expressed genes in a given host cell. Computer algorithms which incorporate codon frequency tables such as "Eco\_high.Cod" for codon preference of highly expressed

bacterial genes may be used and are provided by the University of Wisconsin Package Version 9.0 (Genetics Computer Group, Madison, WI). Other useful codon frequency tables include "Celegans\_high.cod," "Celegans\_low.cod," "Drosophila\_high.cod," "Human\_high.cod," "Maize\_high.cod," and  
5 "Yeast\_high.cod."

In some cases, it may be desirable to prepare nucleic acid molecules encoding ABCL polypeptide variants. Nucleic acid molecules encoding variants may be produced using site directed mutagenesis, PCR amplification, or other appropriate methods, where the primer(s) have the desired point mutations (*see* Sambrook *et al.*,  
10 *supra*, and Ausubel *et al.*, *supra*, for descriptions of mutagenesis techniques). Chemical synthesis using methods described by Engels *et al.*, *supra*, may also be used to prepare such variants. Other methods known to the skilled artisan may be used as well.

#### 15 Vectors and Host Cells

A nucleic acid molecule encoding the amino acid sequence of an ABCL polypeptide is inserted into an appropriate expression vector using standard ligation techniques. The vector is typically selected to be functional in the particular host cell employed (*i.e.*, the vector is compatible with the host cell machinery such that  
20 amplification of the gene and/or expression of the gene can occur). A nucleic acid molecule encoding the amino acid sequence of an ABCL polypeptide may be amplified/expressed in prokaryotic, yeast, insect (baculovirus systems) and/or eukaryotic host cells. Selection of the host cell will depend in part on whether an ABCL polypeptide is to be post-translationally modified (*e.g.*, glycosylated and/or  
25 phosphorylated). If so, yeast, insect, or mammalian host cells are preferable. For a review of expression vectors, *see Meth. Enz.*, vol. 185 (D.V. Goeddel, ed., Academic Press 1990).

Typically, expression vectors used in any of the host cells will contain sequences for plasmid maintenance and for cloning and expression of exogenous  
30 nucleotide sequences. Such sequences, collectively referred to as "flanking sequences" in certain embodiments will typically include one or more of the following nucleotide sequences: a promoter, one or more enhancer sequences, an

origin of replication, a transcriptional termination sequence, a complete intron sequence containing a donor and acceptor splice site, a sequence encoding a leader sequence for polypeptide secretion, a ribosome binding site, a polyadenylation sequence, a polylinker region for inserting the nucleic acid encoding the polypeptide to be expressed, and a selectable marker element. Each of these sequences is discussed below.

Optionally, the vector may contain a “tag”-encoding sequence, *i.e.*, an oligonucleotide molecule located at the 5’ or 3’ end of the ABCL polypeptide coding sequence; the oligonucleotide sequence encodes polyHis (such as hexaHis), or another “tag” such as FLAG, HA (hemagglutinin influenza virus), or *myc* for which commercially available antibodies exist. This tag is typically fused to the polypeptide upon expression of the polypeptide, and can serve as a means for affinity purification of the ABCL polypeptide from the host cell. Affinity purification can be accomplished, for example, by column chromatography using antibodies against the tag as an affinity matrix. Optionally, the tag can subsequently be removed from the purified ABCL polypeptide by various means such as using certain peptidases for cleavage.

Flanking sequences may be homologous (*i.e.*, from the same species and/or strain as the host cell), heterologous (*i.e.*, from a species other than the host cell species or strain), hybrid (*i.e.*, a combination of flanking sequences from more than one source), or synthetic, or the flanking sequences may be native sequences that normally function to regulate ABCL polypeptide expression. As such, the source of a flanking sequence may be any prokaryotic or eukaryotic organism, any vertebrate or invertebrate organism, or any plant, provided that the flanking sequence is functional in, and can be activated by, the host cell machinery.

Flanking sequences useful in the vectors of this invention may be obtained by any of several methods well known in the art. Typically, flanking sequences useful herein – other than the ABCL gene flanking sequences – will have been previously identified by mapping and/or by restriction endonuclease digestion and can thus be isolated from the proper tissue source using the appropriate restriction endonucleases. In some cases, the full nucleotide sequence of a flanking sequence may be known.

Here, the flanking sequence may be synthesized using the methods described herein for nucleic acid synthesis or cloning.

Where all or only a portion of the flanking sequence is known, it may be obtained using PCR and/or by screening a genomic library with a suitable oligonucleotide and/or flanking sequence fragment from the same or another species. Where the flanking sequence is not known, a fragment of DNA containing a flanking sequence may be isolated from a larger piece of DNA that may contain, for example, a coding sequence or even another gene or genes. Isolation may be accomplished by restriction endonuclease digestion to produce the proper DNA fragment followed by isolation using agarose gel purification, Qiagen® column chromatography (Chatsworth, CA), or other methods known to the skilled artisan. The selection of suitable enzymes to accomplish this purpose will be readily apparent to one of ordinary skill in the art.

An origin of replication is typically a part of those prokaryotic expression vectors purchased commercially, and the origin aids in the amplification of the vector in a host cell. Amplification of the vector to a certain copy number can, in some cases, be important for the optimal expression of an ABCL polypeptide. If the vector of choice does not contain an origin of replication site, one may be chemically synthesized based on a known sequence, and ligated into the vector. For example, the origin of replication from the plasmid pBR322 (New England Biolabs, Beverly, MA) is suitable for most gram-negative bacteria and various origins (*e.g.*, SV40, polyoma, adenovirus, vesicular stomatitis virus (VSV), or papillomaviruses such as HPV or BPV) are useful for cloning vectors in mammalian cells. Generally, the origin of replication component is not needed for mammalian expression vectors (for example, the SV40 origin is often used only because it contains the early promoter).

A transcription termination sequence is typically located 3' of the end of a polypeptide coding region and serves to terminate transcription. Usually, a transcription termination sequence in prokaryotic cells is a G-C rich fragment followed by a poly-T sequence. While the sequence is easily cloned from a library or even purchased commercially as part of a vector, it can also be readily synthesized using methods for nucleic acid synthesis such as those described herein.

A selectable marker gene element encodes a protein necessary for the survival and growth of a host cell grown in a selective culture medium. Typical selection marker genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, tetracycline, or kanamycin for prokaryotic host cells; (b) complement auxotrophic deficiencies of the cell; or (c) supply critical nutrients not available from complex media. Preferred selectable markers are the kanamycin resistance gene, the ampicillin resistance gene, and the tetracycline resistance gene. A neomycin resistance gene may also be used for selection in prokaryotic and eukaryotic host cells.

Other selection genes may be used to amplify the gene that will be expressed. Amplification is the process wherein genes that are in greater demand for the production of a protein critical for growth are reiterated in tandem within the chromosomes of successive generations of recombinant cells. Examples of suitable selectable markers for mammalian cells include dihydrofolate reductase (DHFR) and thymidine kinase. The mammalian cell transformants are placed under selection pressure wherein only the transformants are uniquely adapted to survive by virtue of the selection gene present in the vector. Selection pressure is imposed by culturing the transformed cells under conditions in which the concentration of selection agent in the medium is successively changed, thereby leading to the amplification of both the selection gene and the DNA that encodes an ABCL polypeptide. As a result, increased quantities of ABCL polypeptide are synthesized from the amplified DNA.

A ribosome binding site is usually necessary for translation initiation of mRNA and is characterized by a Shine-Dalgarno sequence (prokaryotes) or a Kozak sequence (eukaryotes). The element is typically located 3' to the promoter and 5' to the coding sequence of an ABCL polypeptide to be expressed. The Shine-Dalgarno sequence is varied but is typically a polypurine (*i.e.*, having a high A-G content). Many Shine-Dalgarno sequences have been identified, each of which can be readily synthesized using methods set forth herein and used in a prokaryotic vector.

A leader, or signal, sequence may be used to direct an ABCL polypeptide out of the host cell. Typically, a nucleotide sequence encoding the signal sequence is positioned in the coding region of an ABCL nucleic acid molecule, or directly at the 5' end of an ABCL polypeptide coding region. Many signal sequences have been

identified, and any of those that are functional in the selected host cell may be used in conjunction with an ABCL nucleic acid molecule. Therefore, a signal sequence may be homologous (naturally occurring) or heterologous to the ABCL nucleic acid molecule. Additionally, a signal sequence may be chemically synthesized using  
5 methods described herein. In most cases, the secretion of an ABCL polypeptide from the host cell via the presence of a signal peptide will result in the removal of the signal peptide from the secreted ABCL polypeptide. The signal sequence may be a component of the vector, or it may be a part of an ABCL nucleic acid molecule that is inserted into the vector.

10 Included within the scope of this invention is the use of either a nucleotide sequence encoding a native ABCL polypeptide signal sequence joined to an ABCL polypeptide coding region or a nucleotide sequence encoding a heterologous signal sequence joined to an ABCL polypeptide coding region. The heterologous signal sequence selected should be one that is recognized and processed, *i.e.*, cleaved by a  
15 signal peptidase, by the host cell. For prokaryotic host cells that do not recognize and process the native ABCL polypeptide signal sequence, the signal sequence is substituted by a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, or heat-stable enterotoxin II leaders. For yeast secretion, the native ABCL polypeptide signal sequence may be substituted by the  
20 yeast invertase, alpha factor, or acid phosphatase leaders. In mammalian cell expression the native signal sequence is satisfactory, although other mammalian signal sequences may be suitable.

In some cases, such as where glycosylation is desired in a eukaryotic host cell expression system, one may manipulate the various presequences to improve  
25 glycosylation or yield. For example, one may alter the peptidase cleavage site of a particular signal peptide, or add pro-sequences, which also may affect glycosylation. The final protein product may have, in the -1 position (relative to the first amino acid of the mature protein) one or more additional amino acids incident to expression, which may not have been totally removed. For example, the final protein product  
30 may have one or two amino acid residues found in the peptidase cleavage site, attached to the amino-terminus. Alternatively, use of some enzyme cleavage sites

may result in a slightly truncated form of the desired ABCL polypeptide, if the enzyme cuts at such area within the mature polypeptide.

In many cases, transcription of a nucleic acid molecule is increased by the presence of one or more introns in the vector; this is particularly true where a polypeptide is produced in eukaryotic host cells, especially mammalian host cells. The introns used may be naturally occurring within the ABCL gene especially where the gene used is a full-length genomic sequence or a fragment thereof. Where the intron is not naturally occurring within the gene (as for most cDNAs), the intron may be obtained from another source. The position of the intron with respect to flanking sequences and the ABCL gene is generally important, as the intron must be transcribed to be effective. Thus, when an ABCL cDNA molecule is being transcribed, the preferred position for the intron is 3' to the transcription start site and 5' to the poly-A transcription termination sequence. Preferably, the intron or introns will be located on one side or the other (*i.e.*, 5' or 3') of the cDNA such that it does not interrupt the coding sequence. Any intron from any source, including viral, prokaryotic and eukaryotic (plant or animal) organisms, may be used to practice this invention, provided that it is compatible with the host cell into which it is inserted. Also included herein are synthetic introns. Optionally, more than one intron may be used in the vector.

The expression and cloning vectors of the present invention will typically contain a promoter that is recognized by the host organism and operably linked to the molecule encoding the ABCL polypeptide. Promoters are untranscribed sequences located upstream (*i.e.*, 5') to the start codon of a structural gene (generally within about 100 to 1000 bp) that control the transcription of the structural gene. Promoters are conventionally grouped into one of two classes: inducible promoters and constitutive promoters. Inducible promoters initiate increased levels of transcription from DNA under their control in response to some change in culture conditions, such as the presence or absence of a nutrient or a change in temperature. Constitutive promoters, on the other hand, initiate continual gene product production; that is, there is little or no control over gene expression. A large number of promoters, recognized by a variety of potential host cells, are well known. A suitable promoter is operably linked to the DNA encoding ABCL polypeptide by removing the promoter from the

source DNA by restriction enzyme digestion and inserting the desired promoter sequence into the vector. The native ABCL promoter sequence may be used to direct amplification and/or expression of an ABCL nucleic acid molecule. A heterologous promoter is preferred, however, if it permits greater transcription and higher yields of the expressed protein as compared to the native promoter, and if it is compatible with the host cell system that has been selected for use.

Promoters suitable for use with prokaryotic hosts include the beta-lactamase and lactose promoter systems; alkaline phosphatase; a tryptophan (trp) promoter system; and hybrid promoters such as the tac promoter. Other known bacterial promoters are also suitable. Their sequences have been published, thereby enabling one skilled in the art to ligate them to the desired DNA sequence, using linkers or adapters as needed to supply any useful restriction sites.

Suitable promoters for use with yeast hosts are also well known in the art. Yeast enhancers are advantageously used with yeast promoters. Suitable promoters for use with mammalian host cells are well known and include, but are not limited to, those obtained from the genomes of viruses such as polyoma virus, fowlpox virus, adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, retroviruses, hepatitis-B virus and most preferably Simian Virus 40 (SV40). Other suitable mammalian promoters include heterologous mammalian promoters, for example, heat-shock promoters and the actin promoter.

Additional promoters which may be of interest in controlling ABCL gene expression include, but are not limited to: the SV40 early promoter region (Bernoist and Chambon, 1981, *Nature* 290:304-10); the CMV promoter; the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto, *et al.*, 1980, *Cell* 22:787-97); the herpes thymidine kinase promoter (Wagner *et al.*, 1981, *Proc. Natl. Acad. Sci. U.S.A.* 78:1444-45); the regulatory sequences of the metallothionein gene (Brinster *et al.*, 1982, *Nature* 296:39-42); prokaryotic expression vectors such as the beta-lactamase promoter (Villa-Kamaroff *et al.*, 1978, *Proc. Natl. Acad. Sci. U.S.A.*, 75:3727-31); or the tac promoter (DeBoer *et al.*, 1983, *Proc. Natl. Acad. Sci. U.S.A.*, 80:21-25). Also of interest are the following animal transcriptional control regions, which exhibit tissue specificity and have been utilized in transgenic animals: the elastase I gene control region which is active in pancreatic acinar cells (Swift *et al.*,

1984, *Cell* 38:639-46; Ornitz *et al.*, 1986, *Cold Spring Harbor Symp. Quant. Biol.* 50:399-409 (1986); MacDonald, 1987, *Hepatology* 7:425-515); the insulin gene control region which is active in pancreatic beta cells (Hanahan, 1985, *Nature* 315:115-22); the immunoglobulin gene control region which is active in lymphoid  
5 cells (Grosschedl *et al.*, 1984, *Cell* 38:647-58; Adames *et al.*, 1985, *Nature* 318:533-38; Alexander *et al.*, 1987, *Mol. Cell. Biol.*, 7:1436-44); the mouse mammary tumor virus control region which is active in testicular, breast, lymphoid and mast cells (Leder *et al.*, 1986, *Cell* 45:485-95); the albumin gene control region which is active in liver (Pinkert *et al.*, 1987, *Genes and Devel.* 1:268-76); the alpha-feto-protein gene  
10 control region which is active in liver (Krumlauf *et al.*, 1985, *Mol. Cell. Biol.*, 5:1639-48; Hammer *et al.*, 1987, *Science* 235:53-58); the alpha 1-antitrypsin gene control region which is active in the liver (Kelsey *et al.*, 1987, *Genes and Devel.* 1:161-71); the beta-globin gene control region which is active in myeloid cells (Mogran *et al.*, 1985, *Nature* 315:338-40; Kollias *et al.*, 1986, *Cell* 46:89-94); the myelin basic  
15 protein gene control region which is active in oligodendrocyte cells in the brain (Readhead *et al.*, 1987, *Cell* 48:703-12); the myosin light chain-2 gene control region which is active in skeletal muscle (Sani, 1985, *Nature* 314:283-86); and the gonadotropic releasing hormone gene control region which is active in the hypothalamus (Mason *et al.*, 1986, *Science* 234:1372-78).

20 An enhancer sequence may be inserted into the vector to increase the transcription of a DNA encoding an ABCL polypeptide of the present invention by higher eukaryotes. Enhancers are cis-acting elements of DNA, usually about 10-300 bp in length, that act on the promoter to increase transcription. Enhancers are relatively orientation and position independent. They have been found 5' and 3' to  
25 the transcription unit. Several enhancer sequences available from mammalian genes are known (*e.g.*, globin, elastase, albumin, alpha-feto-protein and insulin). Typically, however, an enhancer from a virus will be used. The SV40 enhancer, the cytomegalovirus early promoter enhancer, the polyoma enhancer, and adenovirus enhancers are exemplary enhancing elements for the activation of eukaryotic  
30 promoters. While an enhancer may be spliced into the vector at a position 5' or 3' to an ABCL nucleic acid molecule, it is typically located at a site 5' from the promoter.

Expression vectors of the invention may be constructed from a starting vector such as a commercially available vector. Such vectors may or may not contain all of the desired flanking sequences. Where one or more of the flanking sequences described herein are not already present in the vector, they may be individually  
5 obtained and ligated into the vector. Methods used for obtaining each of the flanking sequences are well known to one skilled in the art.

Preferred vectors for practicing this invention are those that are compatible with bacterial, insect, and mammalian host cells. Such vectors include, *inter alia*, pCRII, pCR3, and pcDNA3.1 (Invitrogen, San Diego, CA), pBSII (Stratagene, La  
10 Jolla, CA), pET15 (Novagen, Madison, WI), pGEX (Pharmacia Biotech, Piscataway, NJ), pEGFP-N2 (Clontech, Palo Alto, CA), pETL (BlueBacII, Invitrogen), pDSR-alpha (International Pub. No. WO 90/14363) and pFastBacDual (Gibco-BRL, Grand Island, NY).

Additional suitable vectors include, but are not limited to, cosmids, plasmids,  
15 or modified viruses, but it will be appreciated that the vector system must be compatible with the selected host cell. Such vectors include, but are not limited to plasmids such as Bluescript<sup>®</sup> plasmid derivatives (a high copy number ColE1-based phagemid; Stratagene Cloning Systems, La Jolla CA), PCR cloning plasmids designed for cloning Taq-amplified PCR products (*e.g.*, TOPO<sup>™</sup> TA Cloning<sup>®</sup> Kit  
20 and PCR2.1<sup>®</sup> plasmid derivatives; Invitrogen), and mammalian, yeast or virus vectors such as a baculovirus expression system (pBacPAK plasmid derivatives; Clontech).

After the vector has been constructed and a nucleic acid molecule encoding an ABCL polypeptide has been inserted into the proper site of the vector, the completed vector may be inserted into a suitable host cell for amplification and/or polypeptide  
25 expression. The transformation of an expression vector for an ABCL polypeptide into a selected host cell may be accomplished by well known methods including methods such as transfection, infection, calcium chloride, electroporation, microinjection, lipofection, DEAE-dextran method, or other known techniques. The method selected will in part be a function of the type of host cell to be used. These  
30 methods and other suitable methods are well known to the skilled artisan, and are set forth, for example, in Sambrook *et al.*, *supra*.

Host cells may be prokaryotic host cells (such as *E. coli*) or eukaryotic host cells (such as a yeast, insect, or vertebrate cell). The host cell, when cultured under appropriate conditions, synthesizes an ABCL polypeptide that can subsequently be collected from the culture medium (if the host cell secretes it into the medium) or  
5 directly from the host cell producing it (if it is not secreted). The selection of an appropriate host cell will depend upon various factors, such as desired expression levels, polypeptide modifications that are desirable or necessary for activity (such as glycosylation or phosphorylation) and ease of folding into a biologically active molecule.

10 A number of suitable host cells are known in the art and many are available from the American Type Culture Collection (ATCC), Manassas, VA. Examples include, but are not limited to, mammalian cells, such as Chinese hamster ovary cells (CHO), CHO DHFR(-) cells (Urlaub *et al.*, 1980, *Proc. Natl. Acad. Sci. U.S.A.* 97:4216-20), human embryonic kidney (HEK) 293 or 293T cells, or 3T3 cells. The  
15 selection of suitable mammalian host cells and methods for transformation, culture, amplification, screening, product production, and purification are known in the art. Other suitable mammalian cell lines, are the monkey COS-1 and COS-7 cell lines, and the CV-1 cell line. Further exemplary mammalian host cells include primate cell lines and rodent cell lines, including transformed cell lines. Normal diploid cells, cell  
20 strains derived from *in vitro* culture of primary tissue, as well as primary explants, are also suitable. Candidate cells may be genotypically deficient in the selection gene, or may contain a dominantly acting selection gene. Other suitable mammalian cell lines include but are not limited to, mouse neuroblastoma N2A cells, HeLa, mouse L-929 cells, 3T3 lines derived from Swiss, Balb-c or NIH mice, BHK or HaK hamster cell  
25 lines. Each of these cell lines is known by and available to those skilled in the art of protein expression.

Similarly useful as host cells suitable for the present invention are bacterial cells. For example, the various strains of *E. coli* (*e.g.*, HB101, DH5 $\alpha$ , DH10, and MC1061) are well-known as host cells in the field of biotechnology. Various strains  
30 of *B. subtilis*, *Pseudomonas spp.*, other *Bacillus spp.*, *Streptomyces spp.*, and the like may also be employed in this method.

Many strains of yeast cells known to those skilled in the art are also available as host cells for the expression of the polypeptides of the present invention. Preferred yeast cells include, for example, *Saccharomyces cerevisiae* and *Pichia pastoris*.

5 Additionally, where desired, insect cell systems may be utilized in the methods of the present invention. Such systems are described, for example, in Kitts *et al.*, 1993, *Biotechniques*, 14:810-17; Lucklow, 1993, *Curr. Opin. Biotechnol.* 4:564-72; and Lucklow *et al.*, 1993, *J. Virol.*, 67:4566-79. Preferred insect cells are Sf-9 and Hi5 (Invitrogen).

10 One may also use transgenic animals to express glycosylated ABCL polypeptides. For example, one may use a transgenic milk-producing animal (a cow or goat, for example) and obtain the present glycosylated polypeptide in the animal milk. One may also use plants to produce ABCL polypeptides, however, in general, the glycosylation occurring in plants is different from that produced in mammalian cells, and may result in a glycosylated product which is not suitable for human  
15 therapeutic use.

#### Polypeptide Production

Host cells comprising an ABCL polypeptide expression vector may be cultured using standard media well known to the skilled artisan. The media will  
20 usually contain all nutrients necessary for the growth and survival of the cells. Suitable media for culturing *E. coli* cells include, for example, Luria Broth (LB) and/or Terrific Broth (TB). Suitable media for culturing eukaryotic cells include Roswell Park Memorial Institute medium 1640 (RPMI 1640), Minimal Essential Medium (MEM) and/or Dulbecco's Modified Eagle Medium (DMEM), all of which  
25 may be supplemented with serum and/or growth factors as necessary for the particular cell line being cultured. A suitable medium for insect cultures is Grace's medium supplemented with yeastolate, lactalbumin hydrolysate, and/or fetal calf serum as necessary.

Typically, an antibiotic or other compound useful for selective growth of  
30 transfected or transformed cells is added as a supplement to the media. The compound to be used will be dictated by the selectable marker element present on the plasmid with which the host cell was transformed. For example, where the selectable

marker element is kanamycin resistance, the compound added to the culture medium will be kanamycin. Other compounds for selective growth include ampicillin, tetracycline, and neomycin.

5 The amount of an ABCL polypeptide produced by a host cell can be evaluated using standard methods known in the art. Such methods include, without limitation, Western blot analysis, SDS-polyacrylamide gel electrophoresis, non-denaturing gel electrophoresis, High Performance Liquid Chromatography (HPLC) separation, immunoprecipitation, and/or activity assays such as DNA binding gel shift assays.

10 If an ABCL polypeptide has been designed to be secreted from the host cells, the majority of polypeptide may be found in the cell culture medium. If however, the ABCL polypeptide is not secreted from the host cells, it will be present in the cytoplasm and/or the nucleus (for eukaryotic host cells) or in the cytosol (for gram-negative bacteria host cells).

15 For an ABCL polypeptide situated in the host cell cytoplasm and/or nucleus (for eukaryotic host cells) or in the cytosol (for bacterial host cells), the intracellular material (including inclusion bodies for gram-negative bacteria) can be extracted from the host cell using any standard technique known to the skilled artisan. For example, the host cells can be lysed to release the contents of the periplasm/cytoplasm by French press, homogenization, and/or sonication followed by  
20 centrifugation.

If an ABCL polypeptide has formed inclusion bodies in the cytosol, the inclusion bodies can often bind to the inner and/or outer cellular membranes and thus will be found primarily in the pellet material after centrifugation. The pellet material can then be treated at pH extremes or with a chaotropic agent such as a detergent,  
25 guanidine, guanidine derivatives, urea, or urea derivatives in the presence of a reducing agent such as dithiothreitol at alkaline pH or tris carboxyethyl phosphine at acid pH to release, break apart, and solubilize the inclusion bodies. The solubilized ABCL polypeptide can then be analyzed using gel electrophoresis, immunoprecipitation, or the like. If it is desired to isolate the ABCL polypeptide,  
30 isolation may be accomplished using standard methods such as those described herein and in Marston *et al.*, 1990, *Meth. Enz.*, 182:264-75.

In some cases, an ABCL polypeptide may not be biologically active upon isolation. Various methods for "refolding" or converting the polypeptide to its tertiary structure and generating disulfide linkages can be used to restore biological activity. Such methods include exposing the solubilized polypeptide to a pH usually  
5 above 7 and in the presence of a particular concentration of a chaotrope. The selection of chaotrope is very similar to the choices used for inclusion body solubilization, but usually the chaotrope is used at a lower concentration and is not necessarily the same as chaotropes used for the solubilization. In most cases the refolding/oxidation solution will also contain a reducing agent or the reducing agent  
10 plus its oxidized form in a specific ratio to generate a particular redox potential allowing for disulfide shuffling to occur in the formation of the protein's cysteine bridges. Some of the commonly used redox couples include cysteine/cystamine, glutathione (GSH)/dithiobis GSH, cupric chloride, dithiothreitol (DTT)/dithiane DTT, and 2-2-mercaptoethanol (BME)/dithio-BME. In many instances, a cosolvent may  
15 be used or may be needed to increase the efficiency of the refolding, and the more common reagents used for this purpose include glycerol, polyethylene glycol of various molecular weights, arginine and the like.

If inclusion bodies are not formed to a significant degree upon expression of an ABCL polypeptide, then the polypeptide will be found primarily in the supernatant  
20 after centrifugation of the cell homogenate. The polypeptide may be further isolated from the supernatant using methods such as those described herein.

The purification of an ABCL polypeptide from solution can be accomplished using a variety of techniques. If the polypeptide has been synthesized such that it contains a tag such as Hexahistidine (ABCL polypeptide/hexaHis) or other small  
25 peptide such as FLAG (Eastman Kodak Co., New Haven, CT) or *myc* (Invitrogen) at either its carboxyl- or amino-terminus, it may be purified in a one-step process by passing the solution through an affinity column where the column matrix has a high affinity for the tag.

For example, polyhistidine binds with great affinity and specificity to nickel.  
30 Thus, an affinity column of nickel (such as the Qiagen<sup>®</sup> nickel columns) can be used for purification of ABCL polypeptide/polyHis. *See, e.g., Current Protocols in*

*Molecular Biology* § 10.11.8 (Ausubel *et al.*, eds., Green Publishers Inc. and Wiley and Sons 1993).

5 Additionally, ABCL polypeptides may be purified through the use of a monoclonal antibody that is capable of specifically recognizing and binding to an ABCL polypeptide.

Other suitable procedures for purification include, without limitation, affinity chromatography, immunoaffinity chromatography, ion exchange chromatography, molecular sieve chromatography, HPLC, electrophoresis (including native gel electrophoresis) followed by gel elution, and preparative isoelectric focusing  
10 (“Isoprime” machine/technique, Hoefer Scientific, San Francisco, CA). In some cases, two or more purification techniques may be combined to achieve increased purity.

ABCL polypeptides may also be prepared by chemical synthesis methods (such as solid phase peptide synthesis) using techniques known in the art such as  
15 those set forth by Merrifield *et al.*, 1963, *J. Am. Chem. Soc.* 85:2149; Houghten *et al.*, 1985, *Proc Natl Acad. Sci. USA* 82:5132; and Stewart and Young, *Solid Phase Peptide Synthesis* (Pierce Chemical Co. 1984). Such polypeptides may be synthesized with or without a methionine on the amino-terminus. Chemically synthesized ABCL polypeptides may be oxidized using methods set forth in these  
20 references to form disulfide bridges. Chemically synthesized ABCL polypeptides are expected to have comparable biological activity to the corresponding ABCL polypeptides produced recombinantly or purified from natural sources, and thus may be used interchangeably with a recombinant or natural ABCL polypeptide.

Another means of obtaining ABCL polypeptide is via purification from  
25 biological samples such as source tissues and/or fluids in which the ABCL polypeptide is naturally found. Such purification can be conducted using methods for protein purification as described herein. The presence of the ABCL polypeptide during purification may be monitored, for example, using an antibody prepared against recombinantly produced ABCL polypeptide or peptide fragments thereof.

30 A number of additional methods for producing nucleic acids and polypeptides are known in the art, and the methods can be used to produce polypeptides having specificity for ABCL polypeptide. *See, e.g.*, Roberts *et al.*, 1997, *Proc. Natl. Acad.*

*Sci. U.S.A.* 94:12297-303, which describes the production of fusion proteins between an mRNA and its encoded peptide. *See also*, Roberts, 1999, *Curr. Opin. Chem. Biol.* 3:268-73. Additionally, U.S. Patent No. 5,824,469 describes methods for obtaining oligonucleotides capable of carrying out a specific biological function. The procedure involves generating a heterogeneous pool of oligonucleotides, each having a 5' randomized sequence, a central preselected sequence, and a 3' randomized sequence. The resulting heterogeneous pool is introduced into a population of cells that do not exhibit the desired biological function. Subpopulations of the cells are then screened for those that exhibit a predetermined biological function. From that subpopulation, oligonucleotides capable of carrying out the desired biological function are isolated.

U.S. Patent Nos. 5,763,192; 5,814,476; 5,723,323; and 5,817,483 describe processes for producing peptides or polypeptides. This is done by producing stochastic genes or fragments thereof, and then introducing these genes into host cells which produce one or more proteins encoded by the stochastic genes. The host cells are then screened to identify those clones producing peptides or polypeptides having the desired activity.

Another method for producing peptides or polypeptides is described in International Pub. No. WO99/15650, filed by Athersys, Inc. Known as "Random Activation of Gene Expression for Gene Discovery" (RAGE-GD), the process involves the activation of endogenous gene expression or over-expression of a gene by *in situ* recombination methods. For example, expression of an endogenous gene is activated or increased by integrating a regulatory sequence into the target cell that is capable of activating expression of the gene by non-homologous or illegitimate recombination. The target DNA is first subjected to radiation, and a genetic promoter inserted. The promoter eventually locates a break at the front of a gene, initiating transcription of the gene. This results in expression of the desired peptide or polypeptide.

It will be appreciated that these methods can also be used to create comprehensive ABCL polypeptide expression libraries, which can subsequently be used for high throughput phenotypic screening in a variety of assays, such as

biochemical assays, cellular assays, and whole organism assays (*e.g.*, plant, mouse, etc.).

### Synthesis

5           It will be appreciated by those skilled in the art that the nucleic acid and polypeptide molecules described herein may be produced by recombinant and other means.

### Selective Binding Agents

10           The term “selective binding agent” refers to a molecule that has specificity for one or more ABCL polypeptides. Suitable selective binding agents include, but are not limited to, antibodies and derivatives thereof, polypeptides, and small molecules. Suitable selective binding agents may be prepared using methods known in the art. An exemplary ABCL polypeptide selective binding agent of the present invention is  
15           capable of binding a certain portion of the ABCL polypeptide thereby inhibiting the binding of the polypeptide to an ABCL polypeptide receptor.

          Selective binding agents such as antibodies and antibody fragments that bind ABCL polypeptides are within the scope of the present invention. The antibodies may be polyclonal including monospecific polyclonal; monoclonal (MAbs);  
20           recombinant; chimeric; humanized, such as complementarity-determining region (CDR)-grafted; human; single chain; and/or bispecific; as well as fragments; variants; or derivatives thereof. Antibody fragments include those portions of the antibody that bind to an epitope on the ABCL polypeptide. Examples of such fragments include Fab and F(ab') fragments generated by enzymatic cleavage of full-length antibodies.  
25           Other binding fragments include those generated by recombinant DNA techniques, such as the expression of recombinant plasmids containing nucleic acid sequences encoding antibody variable regions.

          Polyclonal antibodies directed toward an ABCL polypeptide generally are produced in animals (*e.g.*, rabbits or mice) by means of multiple subcutaneous or  
30           intraperitoneal injections of ABCL polypeptide and an adjuvant. It may be useful to conjugate an ABCL polypeptide to a carrier protein that is immunogenic in the species to be immunized, such as keyhole limpet hemocyanin, serum, albumin,

bovine thyroglobulin, or soybean trypsin inhibitor. Also, aggregating agents such as alum are used to enhance the immune response. After immunization, the animals are bled and the serum is assayed for anti-ABCL antibody titer.

Monoclonal antibodies directed toward ABCL polypeptides are produced  
5 using any method that provides for the production of antibody molecules by continuous cell lines in culture. Examples of suitable methods for preparing monoclonal antibodies include the hybridoma methods of Kohler *et al.*, 1975, *Nature* 256:495-97 and the human B-cell hybridoma method (Kozbor, 1984, *J. Immunol.* 133:3001; Brodeur *et al.*, *Monoclonal Antibody Production Techniques and*  
10 *Applications* 51-63 (Marcel Dekker, Inc., 1987). Also provided by the invention are hybridoma cell lines that produce monoclonal antibodies reactive with ABCL polypeptides.

Monoclonal antibodies of the invention may be modified for use as therapeutics. One embodiment is a "chimeric" antibody in which a portion of the  
15 heavy (H) and/or light (L) chain is identical with or homologous to a corresponding sequence in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is/are identical with or homologous to a corresponding sequence in antibodies derived from another species or belonging to another antibody class or subclass. Also included are fragments of  
20 such antibodies, so long as they exhibit the desired biological activity. See U.S. Patent No. 4,816,567; Morrison *et al.*, 1985, *Proc. Natl. Acad. Sci.* 81:6851-55.

In another embodiment, a monoclonal antibody of the invention is a "humanized" antibody. Methods for humanizing non-human antibodies are well known in the art. See U.S. Patent Nos. 5,585,089 and 5,693,762. Generally, a  
25 humanized antibody has one or more amino acid residues introduced into it from a source that is non-human. Humanization can be performed, for example, using methods described in the art (Jones *et al.*, 1986, *Nature* 321:522-25; Riechmann *et al.*, 1998, *Nature* 332:323-27; Verhoeyen *et al.*, 1988, *Science* 239:1534-36), by substituting at least a portion of a rodent complementarity-determining region for the  
30 corresponding regions of a human antibody.

Also encompassed by the invention are human antibodies that bind ABCL polypeptides. Using transgenic animals (*e.g.*, mice) that are capable of producing a

repertoire of human antibodies in the absence of endogenous immunoglobulin production such antibodies are produced by immunization with an ABCL polypeptide antigen (*i.e.*, having at least 6 contiguous amino acids), optionally conjugated to a carrier. See, *e.g.*, Jakobovits *et al.*, 1993, *Proc. Natl. Acad. Sci.* 90:2551-55; Jakobovits *et al.*, 1993, *Nature* 362:255-58; Bruggermann *et al.*, 1993, *Year in Immuno.* 7:33. In one method, such transgenic animals are produced by incapacitating the endogenous loci encoding the heavy and light immunoglobulin chains therein, and inserting loci encoding human heavy and light chain proteins into the genome thereof. Partially modified animals (*i.e.*, those having less than the full complement of modifications) are then cross-bred to obtain an animal having all of the desired immune system modifications. When administered an immunogen, these transgenic animals produce antibodies with human (rather than, *e.g.*, murine) amino acid sequences, including variable regions that are immunospecific for these antigens. See International App. Nos. PCT/US96/05928 and PCT/US93/06926. Additional methods are described in U.S. Patent No. 5,545,807, International App. Nos. PCT/US91/245 and PCT/GB89/01207, and in European Patent Nos. 546073B1 and 546073A1. Human antibodies can also be produced by the expression of recombinant DNA in host cells or by expression in hybridoma cells as described herein.

In an alternative embodiment, human antibodies can also be produced from phage-display libraries (Hoogenboom *et al.*, 1991, *J. Mol. Biol.* 227:381; Marks *et al.*, 1991, *J. Mol. Biol.* 222:581). These processes mimic immune selection through the display of antibody repertoires on the surface of filamentous bacteriophage, and subsequent selection of phage by their binding to an antigen of choice. One such technique is described in International App. No. PCT/US98/17364, which describes the isolation of high affinity and functional agonistic antibodies for MPL- and msk-receptors using such an approach.

Chimeric, CDR grafted, and humanized antibodies are typically produced by recombinant methods. Nucleic acids encoding the antibodies are introduced into host cells and expressed using materials and procedures described herein. In a preferred embodiment, the antibodies are produced in mammalian host cells, such as CHO cells. Monoclonal (*e.g.*, human) antibodies may be produced by the expression of

recombinant DNA in host cells or by expression in hybridoma cells as described herein.

The anti-ABCL antibodies of the invention may be employed in any known assay method, such as competitive binding assays, direct and indirect sandwich  
5 assays, and immunoprecipitation assays (Sola, *Monoclonal Antibodies: A Manual of Techniques* 147-158 (CRC Press, Inc., 1987)) for the detection and quantitation of ABCL polypeptides. The antibodies will bind ABCL polypeptides with an affinity that is appropriate for the assay method being employed.

For diagnostic applications, in certain embodiments, anti-ABCL antibodies  
10 may be labeled with a detectable moiety. The detectable moiety can be any one that is capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as  $^3\text{H}$ ,  $^{14}\text{C}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^{125}\text{I}$ ,  $^{99}\text{Tc}$ ,  $^{111}\text{In}$ , or  $^{67}\text{Ga}$ ; a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin; or an enzyme, such as alkaline phosphatase,  
15  $\beta$ -galactosidase, or horseradish peroxidase (Bayer, *et al.*, 1990, *Meth. Enz.* 184:138-63).

Competitive binding assays rely on the ability of a labeled standard (*e.g.*, an ABCL polypeptide, or an immunologically reactive portion thereof) to compete with  
20 the test sample analyte (an ABCL polypeptide) for binding with a limited amount of anti-ABCL antibody. The amount of an ABCL polypeptide in the test sample is inversely proportional to the amount of standard that becomes bound to the antibodies. To facilitate determining the amount of standard that becomes bound, the antibodies typically are insolubilized before or after the competition, so that the standard and analyte that are bound to the antibodies may conveniently be separated  
25 from the standard and analyte that remain unbound.

Sandwich assays typically involve the use of two antibodies, each capable of binding to a different immunogenic portion, or epitope, of the protein to be detected and/or quantitated. In a sandwich assay, the test sample analyte is typically bound by  
30 a first antibody that is immobilized on a solid support, and thereafter a second antibody binds to the analyte, thus forming an insoluble three-part complex. *See, e.g.*, U.S. Patent No. 4,376,110. The second antibody may itself be labeled with a detectable moiety (direct sandwich assays) or may be measured using an anti-

immunoglobulin antibody that is labeled with a detectable moiety (indirect sandwich assays). For example, one type of sandwich assay is an enzyme-linked immunosorbent assay (ELISA), in which case the detectable moiety is an enzyme.

5 The selective binding agents, including anti-ABCL antibodies, are also useful for *in vivo* imaging. An antibody labeled with a detectable moiety may be administered to an animal, preferably into the bloodstream, and the presence and location of the labeled antibody in the host assayed. The antibody may be labeled with any moiety that is detectable in an animal, whether by nuclear magnetic resonance, radiology, or other detection means known in the art.

10 Selective binding agents of the invention, including antibodies, may be used as therapeutics. These therapeutic agents are generally agonists or antagonists, in that they either enhance or reduce, respectively, at least one of the biological activities of an ABCL polypeptide. In one embodiment, antagonist antibodies of the invention are antibodies or binding fragments thereof which are capable of specifically binding to  
15 an ABCL polypeptide and which are capable of inhibiting or eliminating the functional activity of an ABCL polypeptide *in vivo* or *in vitro*. In preferred embodiments, the selective binding agent, *e.g.*, an antagonist antibody, will inhibit the functional activity of an ABCL polypeptide by at least about 50%, and preferably by at least about 80%. In another embodiment, the selective binding agent may be an  
20 anti-ABCL polypeptide antibody that is capable of interacting with an ABCL polypeptide binding partner (a ligand or receptor) thereby inhibiting or eliminating ABCL polypeptide activity *in vitro* or *in vivo*. Selective binding agents, including agonist and antagonist anti-ABCL polypeptide antibodies, are identified by screening assays that are well known in the art.

25 The invention also relates to a kit comprising ABCL selective binding agents (such as antibodies) and other reagents useful for detecting ABCL polypeptide levels in biological samples. Such reagents may include a detectable label, blocking serum, positive and negative control samples, and detection reagents.

### 30 Microarrays

It will be appreciated that DNA microarray technology can be utilized in accordance with the present invention. DNA microarrays are miniature, high-density

arrays of nucleic acids positioned on a solid support, such as glass. Each cell or element within the array contains numerous copies of a single nucleic acid species that acts as a target for hybridization with a complementary nucleic acid sequence (e.g., mRNA). In expression profiling using DNA microarray technology, mRNA is first extracted from a cell or tissue sample and then converted enzymatically to fluorescently labeled cDNA. This material is hybridized to the microarray and unbound cDNA is removed by washing. The expression of discrete genes represented on the array is then visualized by quantitating the amount of labeled cDNA that is specifically bound to each target nucleic acid molecule. In this way, the expression of thousands of genes can be quantitated in a high throughput, parallel manner from a single sample of biological material.

This high throughput expression profiling has a broad range of applications with respect to the ABCL molecules of the invention, including, but not limited to: the identification and validation of ABCL disease-related genes as targets for therapeutics; molecular toxicology of related ABCL molecules and inhibitors thereof; stratification of populations and generation of surrogate markers for clinical trials; and enhancing related ABCL polypeptide small molecule drug discovery by aiding in the identification of selective compounds in high throughput screens.

## Chemical Derivatives

Chemically modified derivatives of ABCL polypeptides may be prepared by one skilled in the art, given the disclosures described herein. ABCL polypeptide derivatives are modified in a manner that is different – either in the type or location of the molecules naturally attached to the polypeptide. Derivatives may include molecules formed by the deletion of one or more naturally-attached chemical groups. The polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, or other ABCL polypeptide, may be modified by the covalent attachment of one or more polymers. For example, the polymer selected is typically water-soluble so that the protein to which it is attached does not precipitate in an aqueous environment, such as a physiological environment. Included within the scope of suitable polymers is a mixture of polymers. Preferably, for therapeutic use of the end-product preparation, the polymer will be pharmaceutically acceptable.

The polymers each may be of any molecular weight and may be branched or unbranched. The polymers each typically have an average molecular weight of between about 2 kDa to about 100 kDa (the term "about" indicating that in preparations of a water-soluble polymer, some molecules will weigh more, some less, than the stated molecular weight). The average molecular weight of each polymer is preferably between about 5 kDa and about 50 kDa, more preferably between about 12 kDa and about 40 kDa and most preferably between about 20 kDa and about 35 kDa.

Suitable water-soluble polymers or mixtures thereof include, but are not limited to, N-linked or O-linked carbohydrates, sugars, phosphates, polyethylene glycol (PEG) (including the forms of PEG that have been used to derivatize proteins, including mono-(C<sub>1</sub>-C<sub>10</sub>), alkoxy-, or aryloxy-polyethylene glycol), monomethoxy-polyethylene glycol, dextran (such as low molecular weight dextran of, for example, about 6 kD), cellulose, or other carbohydrate based polymers, poly-(N-vinyl pyrrolidone) polyethylene glycol, propylene glycol homopolymers, polypropylene oxide/ethylene oxide co-polymers, polyoxyethylated polyols (*e.g.*, glycerol), and polyvinyl alcohol. Also encompassed by the present invention are bifunctional crosslinking molecules that may be used to prepare covalently attached ABCL polypeptide multimers.

In general, chemical derivatization may be performed under any suitable condition used to react a protein with an activated polymer molecule. Methods for preparing chemical derivatives of polypeptides will generally comprise the steps of: (a) reacting the polypeptide with the activated polymer molecule (such as a reactive ester or aldehyde derivative of the polymer molecule) under conditions whereby the polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, or other ABCL polypeptide, becomes attached to one or more polymer molecules, and (b) obtaining the reaction products. The optimal reaction conditions will be determined based on known parameters and the desired result. For example, the larger the ratio of polymer molecules to protein, the greater the percentage of attached polymer molecule. In one embodiment, the ABCL polypeptide derivative may have a single polymer molecule moiety at the amino-terminus. *See, e.g.*, U.S. Patent No. 5,234,784.

The pegylation of a polypeptide may be specifically carried out using any of the pegylation reactions known in the art. Such reactions are described, for example, in the following references: Francis *et al.*, 1992, *Focus on Growth Factors* 3:4-10; European Patent Nos. 0154316 and 0401384; and U.S. Patent No. 4,179,337. For example, pegylation may be carried out via an acylation reaction or an alkylation reaction with a reactive polyethylene glycol molecule (or an analogous reactive water-soluble polymer) as described herein. For the acylation reactions, a selected polymer should have a single reactive ester group. For reductive alkylation, a selected polymer should have a single reactive aldehyde group. A reactive aldehyde is, for example, polyethylene glycol propionaldehyde, which is water stable, or mono C<sub>1</sub>-C<sub>10</sub> alkoxy or aryloxy derivatives thereof (*see* U.S. Patent No. 5,252,714).

In another embodiment, ABCL polypeptides may be chemically coupled to biotin. The biotin/ABCL polypeptide molecules are then allowed to bind to avidin, resulting in tetravalent avidin/biotin/ABCL polypeptide molecules. ABCL polypeptides may also be covalently coupled to dinitrophenol (DNP) or trinitrophenol (TNP) and the resulting conjugates precipitated with anti-DNP or anti-TNP-IgM to form decameric conjugates with a valency of 10.

Generally, conditions that may be alleviated or modulated by the administration of the present ABCL polypeptide derivatives include those described herein for ABCL polypeptides. However, the ABCL polypeptide derivatives disclosed herein may have additional activities, enhanced or reduced biological activity, or other characteristics, such as increased or decreased half-life, as compared to the non-derivatized molecules.

#### 25 Genetically Engineered Non-Human Animals

Additionally included within the scope of the present invention are non-human animals such as mice, rats, or other rodents; rabbits, goats, sheep, or other farm animals, in which the genes encoding native ABCL polypeptide have been disrupted (*i.e.*, “knocked out”) such that the level of expression of ABCL polypeptide is significantly decreased or completely abolished. Such animals may be prepared using techniques and methods such as those described in U.S. Patent No. 5,557,032.

The present invention further includes non-human animals such as mice, rats, or other rodents; rabbits, goats, sheep, or other farm animals, in which either the native form of an ABCL gene for that animal or a heterologous ABCL gene is over-expressed by the animal, thereby creating a “transgenic” animal. Such transgenic  
5 animals may be prepared using well known methods such as those described in U.S. Patent No 5,489,743 and International Pub. No. WO 94/28122.

The present invention further includes non-human animals in which the promoter for one or more of the ABCL polypeptides of the present invention is either activated or inactivated (*e.g.*, by using homologous recombination methods) to alter  
10 the level of expression of one or more of the native ABCL polypeptides.

These non-human animals may be used for drug candidate screening. In such screening, the impact of a drug candidate on the animal may be measured. For example, drug candidates may decrease or increase the expression of the ABCL gene. In certain embodiments, the amount of ABCL polypeptide that is produced may be  
15 measured after the exposure of the animal to the drug candidate. Additionally, in certain embodiments, one may detect the actual impact of the drug candidate on the animal. For example, over-expression of a particular gene may result in, or be associated with, a disease or pathological condition. In such cases, one may test a drug candidate’s ability to decrease expression of the gene or its ability to prevent or  
20 inhibit a pathological condition. In other examples, the production of a particular metabolic product such as a fragment of a polypeptide, may result in, or be associated with, a disease or pathological condition. In such cases, one may test a drug candidate’s ability to decrease the production of such a metabolic product or its ability to prevent or inhibit a pathological condition.

25

#### Assaying for Other Modulators of ABCL Polypeptide Activity

In some situations, it may be desirable to identify molecules that are modulators, *i.e.*, agonists or antagonists, of the activity of ABCL polypeptide. Natural or synthetic molecules that modulate ABCL polypeptide may be identified  
30 using one or more screening assays, such as those described herein. Such molecules may be administered either in an *ex vivo* manner or in an *in vivo* manner by injection, or by oral delivery, implantation device, or the like.

“Test molecule” refers to a molecule that is under evaluation for the ability to modulate (*i.e.*, increase or decrease) the activity of an ABCL polypeptide. Most commonly, a test molecule will interact directly with an ABCL polypeptide. However, it is also contemplated that a test molecule may also modulate ABCL polypeptide activity indirectly, such as by affecting ABCL gene expression, or by binding to an ABCL polypeptide binding partner (*e.g.*, receptor or ligand). In one embodiment, a test molecule will bind to an ABCL polypeptide with an affinity constant of at least about  $10^{-6}$  M, preferably about  $10^{-8}$  M, more preferably about  $10^{-9}$  M, and even more preferably about  $10^{-10}$  M.

10       Methods for identifying compounds that interact with ABCL polypeptides are encompassed by the present invention. In certain embodiments, an ABCL polypeptide is incubated with a test molecule under conditions that permit the interaction of the test molecule with an ABCL polypeptide, and the extent of the interaction is measured. The test molecule can be screened in a substantially purified form or in a crude mixture.

15       In certain embodiments, an ABCL polypeptide agonist or antagonist may be a protein, peptide, carbohydrate, lipid, or small molecular weight molecule that interacts with ABCL polypeptide to regulate its activity. Molecules which regulate ABCL polypeptide expression include nucleic acids which are complementary to nucleic acids encoding an ABCL polypeptide, or are complementary to nucleic acids sequences which direct or control the expression of ABCL polypeptide, and which act as anti-sense regulators of expression.

20       Once a test molecule has been identified as interacting with an ABCL polypeptide, the molecule may be further evaluated for its ability to increase or decrease ABCL polypeptide activity. The measurement of the interaction of a test molecule with ABCL polypeptide may be carried out in several formats, including cell-based binding assays, membrane binding assays, solution-phase assays, and immunoassays. In general, a test molecule is incubated with an ABCL polypeptide for a specified period of time, and ABCL polypeptide activity is determined by one or more assays for measuring biological activity.

30       The interaction of test molecules with ABCL polypeptides may also be assayed directly using polyclonal or monoclonal antibodies in an immunoassay.

Alternatively, modified forms of ABCL polypeptides containing epitope tags as described herein may be used in solution and immunoassays.

In the event that ABCL polypeptides display biological activity through an interaction with a binding partner (*e.g.*, a receptor or a ligand), a variety of *in vitro* assays may be used to measure the binding of an ABCL polypeptide to the corresponding binding partner (such as a selective binding agent, receptor, or ligand). These assays may be used to screen test molecules for their ability to increase or decrease the rate and/or the extent of binding of an ABCL polypeptide to its binding partner. In one assay, an ABCL polypeptide is immobilized in the wells of a microtiter plate. Radiolabeled ABCL polypeptide binding partner (for example, iodinated ABCL polypeptide binding partner) and a test molecule can then be added either one at a time (in either order) or simultaneously to the wells. After incubation, the wells can be washed and counted for radioactivity, using a scintillation counter, to determine the extent to which the binding partner bound to the ABCL polypeptide. Typically, a molecule will be tested over a range of concentrations, and a series of control wells lacking one or more elements of the test assays can be used for accuracy in the evaluation of the results. An alternative to this method involves reversing the “positions” of the proteins, *i.e.*, immobilizing ABCL polypeptide binding partner to the microtiter plate wells, incubating with the test molecule and radiolabeled ABCL polypeptide, and determining the extent of ABCL polypeptide binding. *See, e.g., Current Protocols in Molecular Biology*, chap. 18 (Ausubel *et al.*, eds., Green Publishers Inc. and Wiley and Sons 1995).

As an alternative to radiolabeling, an ABCL polypeptide or its binding partner may be conjugated to biotin, and the presence of biotinylated protein can then be detected using streptavidin linked to an enzyme, such as horse radish peroxidase (HRP) or alkaline phosphatase (AP), which can be detected colorimetrically, or by fluorescent tagging of streptavidin. An antibody directed to an ABCL polypeptide or to an ABCL polypeptide binding partner, and which is conjugated to biotin, may also be used for purposes of detection following incubation of the complex with enzyme-linked streptavidin linked to AP or HRP.

An ABCL polypeptide or an ABCL polypeptide binding partner can also be immobilized by attachment to agarose beads, acrylic beads, or other types of such

inert solid phase substrates. The substrate-protein complex can be placed in a solution containing the complementary protein and the test compound. After incubation, the beads can be precipitated by centrifugation, and the amount of binding between an ABCL polypeptide and its binding partner can be assessed using the methods described herein. Alternatively, the substrate-protein complex can be immobilized in a column with the test molecule and complementary protein passing through the column. The formation of a complex between an ABCL polypeptide and its binding partner can then be assessed using any of the techniques described herein (e.g., radiolabelling or antibody binding).

Another *in vitro* assay that is useful for identifying a test molecule that increases or decreases the formation of a complex between an ABCL polypeptide binding protein and an ABCL polypeptide binding partner is a surface plasmon resonance detector system such as the BIAcore assay system (Pharmacia, Piscataway, NJ). The BIAcore system is utilized as specified by the manufacturer. This assay essentially involves the covalent binding of either ABCL polypeptide or an ABCL polypeptide binding partner to a dextran-coated sensor chip that is located in a detector. The test compound and the other complementary protein can then be injected, either simultaneously or sequentially, into the chamber containing the sensor chip. The amount of complementary protein that binds can be assessed based on the change in molecular mass that is physically associated with the dextran-coated side of the sensor chip, with the change in molecular mass being measured by the detector system.

In some cases, it may be desirable to evaluate two or more test compounds together for their ability to increase or decrease the formation of a complex between an ABCL polypeptide and an ABCL polypeptide binding partner. In these cases, the assays set forth herein can be readily modified by adding such additional test compound(s) either simultaneously with, or subsequent to, the first test compound. The remainder of the steps in the assay are as set forth herein.

*In vitro* assays such as those described herein may be used advantageously to screen large numbers of compounds for an effect on the formation of a complex between an ABCL polypeptide and ABCL polypeptide binding partner. The assays

may be automated to screen compounds generated in phage display, synthetic peptide, and chemical synthesis libraries.

Compounds which increase or decrease the formation of a complex between an ABCL polypeptide and an ABCL polypeptide binding partner may also be screened in cell culture using cells and cell lines expressing either ABCL polypeptide or ABCL polypeptide binding partner. Cells and cell lines may be obtained from any mammal, but preferably will be from human or other primate, canine, or rodent sources. The binding of an ABCL polypeptide to cells expressing ABCL polypeptide binding partner at the surface is evaluated in the presence or absence of test molecules, and the extent of binding may be determined by, for example, flow cytometry using a biotinylated antibody to an ABCL polypeptide binding partner. Cell culture assays can be used advantageously to further evaluate compounds that score positive in protein binding assays described herein.

Cell cultures can also be used to screen the impact of a drug candidate. For example, drug candidates may decrease or increase the expression of the ABCL gene. In certain embodiments, the amount of ABCL polypeptide or an ABCL polypeptide fragment that is produced may be measured after exposure of the cell culture to the drug candidate. In certain embodiments, one may detect the actual impact of the drug candidate on the cell culture. For example, the over-expression of a particular gene may have a particular impact on the cell culture. In such cases, one may test a drug candidate's ability to increase or decrease the expression of the gene or its ability to prevent or inhibit a particular impact on the cell culture. In other examples, the production of a particular metabolic product such as a fragment of a polypeptide, may result in, or be associated with, a disease or pathological condition. In such cases, one may test a drug candidate's ability to decrease the production of such a metabolic product in a cell culture.

#### Internalizing Proteins

The *tat* protein sequence (from HIV) can be used to internalize proteins into a cell. See, e.g., Falwell *et al.*, 1994, *Proc. Natl. Acad. Sci. U.S.A.* 91:664-68. For example, an 11 amino acid sequence (Y-G-R-K-K-R-R-Q-R-R-R; SEQ ID NO: 13) of the HIV *tat* protein (termed the "protein transduction domain," or TAT PDT) has

been described as mediating delivery across the cytoplasmic membrane and the nuclear membrane of a cell. See Schwarze *et al.*, 1999, *Science* 285:1569-72; and Nagahara *et al.*, 1998, *Nat. Med.* 4:1449-52. In these procedures, FITC-constructs (FITC-labeled G-G-G-G-Y-G-R-K-K-R-R-Q-R-R-R; SEQ ID NO: 14), which  
5 penetrate tissues following intraperitoneal administration, are prepared, and the binding of such constructs to cells is detected by fluorescence-activated cell sorting (FACS) analysis. Cells treated with a *tat*- $\beta$ -gal fusion protein will demonstrate  $\beta$ -gal activity. Following injection, expression of such a construct can be detected in a number of tissues, including liver, kidney, lung, heart, and brain tissue. It is believed  
10 that such constructs undergo some degree of unfolding in order to enter the cell, and as such, may require a refolding following entry into the cell.

It will thus be appreciated that the *tat* protein sequence may be used to internalize a desired polypeptide into a cell. For example, using the *tat* protein sequence, an ABCL antagonist (such as an anti-ABCL selective binding agent, small  
15 molecule, soluble receptor, or antisense oligonucleotide) can be administered intracellularly to inhibit the activity of an ABCL molecule. As used herein, the term "ABCL molecule" refers to both ABCL nucleic acid molecules and ABCL polypeptides as defined herein. Where desired, the ABCL protein itself may also be internally administered to a cell using these procedures. See also, Straus, 1999,  
20 *Science* 285:1466-67.

#### Cell Source Identification Using ABCL Polypeptide

In accordance with certain embodiments of the invention, it may be useful to be able to determine the source of a certain cell type associated with an ABCL  
25 polypeptide. For example, it may be useful to determine the origin of a disease or pathological condition as an aid in selecting an appropriate therapy. In certain embodiments, nucleic acids encoding an ABCL polypeptide can be used as a probe to identify cells described herein by screening the nucleic acids of the cells with such a probe. In other embodiments, one may use anti-ABCL polypeptide antibodies to test  
30 for the presence of ABCL polypeptide in cells, and thus, determine if such cells are of the types described herein.

ABCL Polypeptide Compositions and Administration

Therapeutic compositions are within the scope of the present invention. Such ABCL polypeptide pharmaceutical compositions may comprise a therapeutically effective amount of an ABCL polypeptide or an ABCL nucleic acid molecule in admixture with a pharmaceutically or physiologically acceptable formulation agent selected for suitability with the mode of administration. Pharmaceutical compositions may comprise a therapeutically effective amount of one or more ABCL polypeptide selective binding agents in admixture with a pharmaceutically or physiologically acceptable formulation agent selected for suitability with the mode of administration.

Acceptable formulation materials preferably are nontoxic to recipients at the dosages and concentrations employed.

The pharmaceutical composition may contain formulation materials for modifying, maintaining, or preserving, for example, the pH, osmolarity, viscosity, clarity, color, isotonicity, odor, sterility, stability, rate of dissolution or release, adsorption, or penetration of the composition. Suitable formulation materials include, but are not limited to, amino acids (such as glycine, glutamine, asparagine, arginine, or lysine), antimicrobials, antioxidants (such as ascorbic acid, sodium sulfite, or sodium hydrogen-sulfite), buffers (such as borate, bicarbonate, Tris-HCl, citrates, phosphates, or other organic acids), bulking agents (such as mannitol or glycine), chelating agents (such as ethylenediamine tetraacetic acid (EDTA)), complexing agents (such as caffeine, polyvinylpyrrolidone, beta-cyclodextrin, or hydroxypropyl-beta-cyclodextrin), fillers, monosaccharides, disaccharides, and other carbohydrates (such as glucose, mannose, or dextrans), proteins (such as serum albumin, gelatin, or immunoglobulins), coloring, flavoring and diluting agents, emulsifying agents, hydrophilic polymers (such as polyvinylpyrrolidone), low molecular weight polypeptides, salt-forming counterions (such as sodium), preservatives (such as benzalkonium chloride, benzoic acid, salicylic acid, thimerosal, phenethyl alcohol, methylparaben, propylparaben, chlorhexidine, sorbic acid, or hydrogen peroxide), solvents (such as glycerin, propylene glycol, or polyethylene glycol), sugar alcohols (such as mannitol or sorbitol), suspending agents, surfactants or wetting agents (such as pluronics; PEG; sorbitan esters; polysorbates such as polysorbate 20 or polysorbate 80; triton; tromethamine; lecithin; cholesterol or tyloxapal), stability enhancing agents

(such as sucrose or sorbitol), tonicity enhancing agents (such as alkali metal halides – preferably sodium or potassium chloride – or mannitol sorbitol), delivery vehicles, diluents, excipients and/or pharmaceutical adjuvants. See *Remington's Pharmaceutical Sciences* (18th Ed., A.R. Gennaro, ed., Mack Publishing Company  
5 1990.

The optimal pharmaceutical composition will be determined by a skilled artisan depending upon, for example, the intended route of administration, delivery format, and desired dosage. See, e.g., *Remington's Pharmaceutical Sciences, supra*. Such compositions may influence the physical state, stability, rate of *in vivo* release,  
10 and rate of *in vivo* clearance of the ABCL molecule.

The primary vehicle or carrier in a pharmaceutical composition may be either aqueous or non-aqueous in nature. For example, a suitable vehicle or carrier for injection may be water, physiological saline solution, or artificial cerebrospinal fluid, possibly supplemented with other materials common in compositions for parenteral  
15 administration. Neutral buffered saline or saline mixed with serum albumin are further exemplary vehicles. Other exemplary pharmaceutical compositions comprise Tris buffer of about pH 7.0-8.5, or acetate buffer of about pH 4.0-5.5, which may further include sorbitol or a suitable substitute. In one embodiment of the present invention, ABCL polypeptide compositions may be prepared for storage by mixing  
20 the selected composition having the desired degree of purity with optional formulation agents (*Remington's Pharmaceutical Sciences, supra*) in the form of a lyophilized cake or an aqueous solution. Further, the ABCL polypeptide product may be formulated as a lyophilizate using appropriate excipients such as sucrose.

The ABCL polypeptide pharmaceutical compositions can be selected for  
25 parenteral delivery. Alternatively, the compositions may be selected for inhalation or for delivery through the digestive tract, such as orally. The preparation of such pharmaceutically acceptable compositions is within the skill of the art.

The formulation components are present in concentrations that are acceptable to the site of administration. For example, buffers are used to maintain the  
30 composition at physiological pH or at a slightly lower pH, typically within a pH range of from about 5 to about 8.

When parenteral administration is contemplated, the therapeutic compositions for use in this invention may be in the form of a pyrogen-free, parenterally acceptable, aqueous solution comprising the desired ABCL molecule in a pharmaceutically acceptable vehicle. A particularly suitable vehicle for parenteral  
5 injection is sterile distilled water in which an ABCL molecule is formulated as a sterile, isotonic solution, properly preserved. Yet another preparation can involve the formulation of the desired molecule with an agent, such as injectable microspheres, bio-erodible particles, polymeric compounds (such as polylactic acid or polyglycolic acid), beads, or liposomes, that provides for the controlled or sustained release of the  
10 product which may then be delivered via a depot injection. Hyaluronic acid may also be used, and this may have the effect of promoting sustained duration in the circulation. Other suitable means for the introduction of the desired molecule include implantable drug delivery devices.

In one embodiment, a pharmaceutical composition may be formulated for  
15 inhalation. For example, ABCL polypeptide may be formulated as a dry powder for inhalation. ABCL polypeptide or nucleic acid molecule inhalation solutions may also be formulated with a propellant for aerosol delivery. In yet another embodiment, solutions may be nebulized. Pulmonary administration is further described in International Pub. No. WO 94/20069, which describes the pulmonary delivery of  
20 chemically modified proteins.

It is also contemplated that certain formulations may be administered orally. In one embodiment of the present invention, ABCL polypeptides that are administered in this fashion can be formulated with or without those carriers customarily used in the compounding of solid dosage forms such as tablets and  
25 capsules. For example, a capsule may be designed to release the active portion of the formulation at the point in the gastrointestinal tract when bioavailability is maximized and pre-systemic degradation is minimized. Additional agents can be included to facilitate absorption of the ABCL polypeptide. Diluents, flavorings, low melting point waxes, vegetable oils, lubricants, suspending agents, tablet disintegrating  
30 agents, and binders may also be employed.

Another pharmaceutical composition may involve an effective quantity of ABCL polypeptides in a mixture with non-toxic excipients that are suitable for the

manufacture of tablets. By dissolving the tablets in sterile water, or another appropriate vehicle, solutions can be prepared in unit-dose form. Suitable excipients include, but are not limited to, inert diluents, such as calcium carbonate, sodium carbonate or bicarbonate, lactose, or calcium phosphate; or binding agents, such as starch, gelatin, or acacia; or lubricating agents such as magnesium stearate, stearic acid, or talc.

Additional ABCL polypeptide pharmaceutical compositions will be evident to those skilled in the art, including formulations involving ABCL polypeptides in sustained- or controlled-delivery formulations. Techniques for formulating a variety of other sustained- or controlled-delivery means, such as liposome carriers, bio-erodible microparticles or porous beads and depot injections, are also known to those skilled in the art. *See, e.g.*, International App. No. PCT/US93/00829, which describes the controlled release of porous polymeric microparticles for the delivery of pharmaceutical compositions.

Additional examples of sustained-release preparations include semipermeable polymer matrices in the form of shaped articles, *e.g.* films, or microcapsules. Sustained release matrices may include polyesters, hydrogels, polylactides (U.S. Patent No. 3,773,919 and European Patent No. 058481), copolymers of L-glutamic acid and gamma ethyl-L-glutamate (Sidman *et al.*, 1983, *Biopolymers* 22:547-56), poly(2-hydroxyethyl-methacrylate) (Langer *et al.*, 1981, *J. Biomed. Mater. Res.* 15:167-277 and Langer, 1982, *Chem. Tech.* 12:98-105), ethylene vinyl acetate (Langer *et al.*, *supra*) or poly-D(-)-3-hydroxybutyric acid (European Patent No. 133988). Sustained-release compositions may also include liposomes, which can be prepared by any of several methods known in the art. *See, e.g.*, Eppstein *et al.*, 1985, *Proc. Natl. Acad. Sci. USA* 82:3688-92; and European Patent Nos. 036676, 088046, and 143949.

The ABCL pharmaceutical composition to be used for *in vivo* administration typically must be sterile. This may be accomplished by filtration through sterile filtration membranes. Where the composition is lyophilized, sterilization using this method may be conducted either prior to, or following, lyophilization and reconstitution. The composition for parenteral administration may be stored in lyophilized form or in a solution. In addition, parenteral compositions generally are

placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Once the pharmaceutical composition has been formulated, it may be stored in sterile vials as a solution, suspension, gel, emulsion, solid, or as a dehydrated or lyophilized powder. Such formulations may be stored either in a ready-to-use form or  
5 in a form (*e.g.*, lyophilized) requiring reconstitution prior to administration.

In a specific embodiment, the present invention is directed to kits for producing a single-dose administration unit. The kits may each contain both a first container having a dried protein and a second container having an aqueous  
10 formulation. Also included within the scope of this invention are kits containing single and multi-chambered pre-filled syringes (*e.g.*, liquid syringes and lyosyringes).

The effective amount of an ABCL pharmaceutical composition to be employed therapeutically will depend, for example, upon the therapeutic context and objectives. One skilled in the art will appreciate that the appropriate dosage levels for  
15 treatment will thus vary depending, in part, upon the molecule delivered, the indication for which the ABCL molecule is being used, the route of administration, and the size (body weight, body surface, or organ size) and condition (the age and general health) of the patient. Accordingly, the clinician may titer the dosage and modify the route of administration to obtain the optimal therapeutic effect. A typical  
20 dosage may range from about 0.1 µg/kg to up to about 100 mg/kg or more, depending on the factors mentioned above. In other embodiments, the dosage may range from 0.1 µg/kg up to about 100 mg/kg; or 1 µg/kg up to about 100 mg/kg; or 5 µg/kg up to about 100 mg/kg.

The frequency of dosing will depend upon the pharmacokinetic parameters of  
25 the ABCL molecule in the formulation being used. Typically, a clinician will administer the composition until a dosage is reached that achieves the desired effect. The composition may therefore be administered as a single dose, as two or more doses (which may or may not contain the same amount of the desired molecule) over time, or as a continuous infusion via an implantation device or catheter. Further  
30 refinement of the appropriate dosage is routinely made by those of ordinary skill in the art and is within the ambit of tasks routinely performed by them. Appropriate dosages may be ascertained through use of appropriate dose-response data.

The route of administration of the pharmaceutical composition is in accord with known methods, *e.g.*, orally; through injection by intravenous, intraperitoneal, intracerebral (intraparenchymal), intracerebroventricular, intramuscular, intraocular, intraarterial, intraportal, or intralesional routes; by sustained release systems; or by  
5 implantation devices. Where desired, the compositions may be administered by bolus injection or continuously by infusion, or by implantation device.

Alternatively or additionally, the composition may be administered locally via implantation of a membrane, sponge, or other appropriate material onto which the desired molecule has been absorbed or encapsulated. Where an implantation device  
10 is used, the device may be implanted into any suitable tissue or organ, and delivery of the desired molecule may be via diffusion, timed-release bolus, or continuous administration.

In some cases, it may be desirable to use ABCL polypeptide pharmaceutical compositions in an *ex vivo* manner. In such instances, cells, tissues, or organs that  
15 have been removed from the patient are exposed to ABCL polypeptide pharmaceutical compositions after which the cells, tissues, or organs are subsequently implanted back into the patient.

In other cases, an ABCL polypeptide can be delivered by implanting certain cells that have been genetically engineered, using methods such as those described  
20 herein, to express and secrete the ABCL polypeptide. Such cells may be animal or human cells, and may be autologous, heterologous, or xenogeneic. Optionally, the cells may be immortalized. In order to decrease the chance of an immunological response, the cells may be encapsulated to avoid infiltration of surrounding tissues. The encapsulation materials are typically biocompatible, semi-permeable polymeric  
25 enclosures or membranes that allow the release of the protein product(s) but prevent the destruction of the cells by the patient's immune system or by other detrimental factors from the surrounding tissues.

As discussed herein, it may be desirable to treat isolated cell populations (such as stem cells, lymphocytes, red blood cells, chondrocytes, neurons, and the like) with  
30 one or more ABCL polypeptides. This can be accomplished by exposing the isolated cells to the polypeptide directly, where it is in a form that is permeable to the cell membrane.

Additional embodiments of the present invention relate to cells and methods (e.g., homologous recombination and/or other recombinant production methods) for both the *in vitro* production of therapeutic polypeptides and for the production and delivery of therapeutic polypeptides by gene therapy or cell therapy. Homologous and other recombination methods may be used to modify a cell that contains a normally transcriptionally-silent ABCL gene, or an under-expressed gene, and thereby produce a cell which expresses therapeutically efficacious amounts of ABCL polypeptides.

Homologous recombination is a technique originally developed for targeting genes to induce or correct mutations in transcriptionally active genes. Kucherlapati, 1989, *Prog. in Nucl. Acid Res. & Mol. Biol.* 36:301. The basic technique was developed as a method for introducing specific mutations into specific regions of the mammalian genome (Thomas *et al.*, 1986, *Cell* 44:419-28; Thomas and Capecchi, 1987, *Cell* 51:503-12; Doetschman *et al.*, 1988, *Proc. Natl. Acad. Sci. U.S.A.* 85:8583-87) or to correct specific mutations within defective genes (Doetschman *et al.*, 1987, *Nature* 330:576-78). Exemplary homologous recombination techniques are described in U.S. Patent No. 5,272,071; European Patent Nos. 9193051 and 505500; International App. No. PCT/US90/07642, and International Pub No. WO 91/09955).

Through homologous recombination, the DNA sequence to be inserted into the genome can be directed to a specific region of the gene of interest by attaching it to targeting DNA. The targeting DNA is a nucleotide sequence that is complementary (homologous) to a region of the genomic DNA. Small pieces of targeting DNA that are complementary to a specific region of the genome are put in contact with the parental strand during the DNA replication process. It is a general property of DNA that has been inserted into a cell to hybridize, and therefore, recombine with other pieces of endogenous DNA through shared homologous regions. If this complementary strand is attached to an oligonucleotide that contains a mutation or a different sequence or an additional nucleotide, it too is incorporated into the newly synthesized strand as a result of the recombination. As a result of the proofreading function, it is possible for the new sequence of DNA to serve as the template. Thus, the transferred DNA is incorporated into the genome.

Attached to these pieces of targeting DNA are regions of DNA that may interact with or control the expression of an ABCL polypeptide, *e.g.*, flanking sequences. For example, a promoter/enhancer element, a suppressor, or an exogenous transcription modulatory element is inserted in the genome of the intended host cell in  
5 proximity and orientation sufficient to influence the transcription of DNA encoding the desired ABCL polypeptide. The control element controls a portion of the DNA present in the host cell genome. Thus, the expression of the desired ABCL polypeptide may be achieved not by transfection of DNA that encodes the ABCL gene itself, but rather by the use of targeting DNA (containing regions of homology  
10 with the endogenous gene of interest) coupled with DNA regulatory segments that provide the endogenous gene sequence with recognizable signals for transcription of an ABCL gene.

In an exemplary method, the expression of a desired targeted gene in a cell (*i.e.*, a desired endogenous cellular gene) is altered via homologous recombination  
15 into the cellular genome at a preselected site, by the introduction of DNA that includes at least a regulatory sequence, an exon, and a splice donor site. These components are introduced into the chromosomal (genomic) DNA in such a manner that this, in effect, results in the production of a new transcription unit (in which the regulatory sequence, the exon, and the splice donor site present in the DNA construct  
20 are operatively linked to the endogenous gene). As a result of the introduction of these components into the chromosomal DNA, the expression of the desired endogenous gene is altered.

Altered gene expression, as described herein, encompasses activating (or causing to be expressed) a gene which is normally silent (unexpressed) in the cell as  
25 obtained, as well as increasing the expression of a gene which is not expressed at physiologically significant levels in the cell as obtained. The embodiments further encompass changing the pattern of regulation or induction such that it is different from the pattern of regulation or induction that occurs in the cell as obtained, and reducing (including eliminating) the expression of a gene which is expressed in the  
30 cell as obtained.

One method by which homologous recombination can be used to increase, or cause, ABCL polypeptide production from a cell's endogenous ABCL gene involves

first using homologous recombination to place a recombination sequence from a site-specific recombination system (*e.g.*, Cre/loxP, FLP/FRT) (Sauer, 1994, *Curr. Opin. Biotechnol.*, 5:521-27; Sauer, 1993, *Methods Enzymol.*, 225:890-900) upstream of (*i.e.*, 5' to) the cell's endogenous genomic ABCL polypeptide coding region. A plasmid containing a recombination site homologous to the site that was placed just upstream of the genomic ABCL polypeptide coding region is introduced into the modified cell line along with the appropriate recombinase enzyme. This recombinase causes the plasmid to integrate, via the plasmid's recombination site, into the recombination site located just upstream of the genomic ABCL polypeptide coding region in the cell line (Baubonis and Sauer, 1993, *Nucleic Acids Res.* 21:2025-29; O'Gorman *et al.*, 1991, *Science* 251:1351-55). Any flanking sequences known to increase transcription (*e.g.*, enhancer/promoter, intron, translational enhancer), if properly positioned in this plasmid, would integrate in such a manner as to create a new or modified transcriptional unit resulting in *de novo* or increased ABCL polypeptide production from the cell's endogenous ABCL gene.

A further method to use the cell line in which the site specific recombination sequence had been placed just upstream of the cell's endogenous genomic ABCL polypeptide coding region is to use homologous recombination to introduce a second recombination site elsewhere in the cell line's genome. The appropriate recombinase enzyme is then introduced into the two-recombination-site cell line, causing a recombination event (deletion, inversion, and translocation) (Sauer, 1994, *Curr. Opin. Biotechnol.*, 5:521-27; Sauer, 1993, *Methods Enzymol.*, 225:890-900) that would create a new or modified transcriptional unit resulting in *de novo* or increased ABCL polypeptide production from the cell's endogenous ABCL gene.

An additional approach for increasing, or causing, the expression of ABCL polypeptide from a cell's endogenous ABCL gene involves increasing, or causing, the expression of a gene or genes (*e.g.*, transcription factors) and/or decreasing the expression of a gene or genes (*e.g.*, transcriptional repressors) in a manner which results in *de novo* or increased ABCL polypeptide production from the cell's endogenous ABCL gene. This method includes the introduction of a non-naturally occurring polypeptide (*e.g.*, a polypeptide comprising a site specific DNA binding domain fused to a transcriptional factor domain) into the cell such that *de novo* or

increased ABCL polypeptide production from the cell's endogenous ABCL gene results.

The present invention further relates to DNA constructs useful in the method of altering expression of a target gene. In certain embodiments, the exemplary DNA constructs comprise: (a) one or more targeting sequences, (b) a regulatory sequence, (c) an exon, and (d) an unpaired splice-donor site. The targeting sequence in the DNA construct directs the integration of elements (a) - (d) into a target gene in a cell such that the elements (b) - (d) are operatively linked to sequences of the endogenous target gene. In another embodiment, the DNA constructs comprise: (a) one or more targeting sequences, (b) a regulatory sequence, (c) an exon, (d) a splice-donor site, (e) an intron, and (f) a splice-acceptor site, wherein the targeting sequence directs the integration of elements (a) - (f) such that the elements of (b) - (f) are operatively linked to the endogenous gene. The targeting sequence is homologous to the preselected site in the cellular chromosomal DNA with which homologous recombination is to occur. In the construct, the exon is generally 3' of the regulatory sequence and the splice-donor site is 3' of the exon.

If the sequence of a particular gene is known, such as the nucleic acid sequence of ABCL polypeptide presented herein, a piece of DNA that is complementary to a selected region of the gene can be synthesized or otherwise obtained, such as by appropriate restriction of the native DNA at specific recognition sites bounding the region of interest. This piece serves as a targeting sequence upon insertion into the cell and will hybridize to its homologous region within the genome. If this hybridization occurs during DNA replication, this piece of DNA, and any additional sequence attached thereto, will act as an Okazaki fragment and will be incorporated into the newly synthesized daughter strand of DNA. The present invention, therefore, includes nucleotides encoding an ABCL polypeptide, which nucleotides may be used as targeting sequences.

ABCL polypeptide cell therapy, *e.g.*, the implantation of cells producing ABCL polypeptides, is also contemplated. This embodiment involves implanting cells capable of synthesizing and secreting a biologically active form of ABCL polypeptide. Such ABCL polypeptide-producing cells can be cells that are natural producers of ABCL polypeptides or may be recombinant cells whose ability to

produce ABCL polypeptides has been augmented by transformation with a gene encoding the desired ABCL polypeptide or with a gene augmenting the expression of ABCL polypeptide. Such a modification may be accomplished by means of a vector suitable for delivering the gene as well as promoting its expression and secretion. In order to minimize a potential immunological reaction in patients being administered an ABCL polypeptide, as may occur with the administration of a polypeptide of a foreign species, it is preferred that the natural cells producing ABCL polypeptide be of human origin and produce human ABCL polypeptide. Likewise, it is preferred that the recombinant cells producing ABCL polypeptide be transformed with an expression vector containing a gene encoding a human ABCL polypeptide.

Implanted cells may be encapsulated to avoid the infiltration of surrounding tissue. Human or non-human animal cells may be implanted in patients in biocompatible, semipermeable polymeric enclosures or membranes that allow the release of ABCL polypeptide, but that prevent the destruction of the cells by the patient's immune system or by other detrimental factors from the surrounding tissue. Alternatively, the patient's own cells, transformed to produce ABCL polypeptides *ex vivo*, may be implanted directly into the patient without such encapsulation.

Techniques for the encapsulation of living cells are known in the art, and the preparation of the encapsulated cells and their implantation in patients may be routinely accomplished. For example, Baetge *et al.* (International Pub. No. WO 95/05452 and International App. No. PCT/US94/09299) describe membrane capsules containing genetically engineered cells for the effective delivery of biologically active molecules. The capsules are biocompatible and are easily retrievable. The capsules encapsulate cells transfected with recombinant DNA molecules comprising DNA sequences coding for biologically active molecules operatively linked to promoters that are not subject to down-regulation *in vivo* upon implantation into a mammalian host. The devices provide for the delivery of the molecules from living cells to specific sites within a recipient. In addition, *see* U.S. Patent Nos. 4,892,538; 5,011,472; and 5,106,627. A system for encapsulating living cells is described in International Pub. No. WO 91/10425 (Aebischer *et al.*). *See also*, International Pub. No. WO 91/10470 (Aebischer *et al.*); Winn *et al.*, 1991, *Exper. Neurol.* 113:322-29;

Aebischer *et al.*, 1991, *Exper. Neurol.* 111:269-75; and Tresco *et al.*, 1992, *ASAIO* 38:17-23.

*In vivo* and *in vitro* gene therapy delivery of ABCL polypeptides is also envisioned. One example of a gene therapy technique is to use the ABCL gene  
5 (either genomic DNA, cDNA, and/or synthetic DNA) encoding an ABCL polypeptide that may be operably linked to a constitutive or inducible promoter to form a "gene therapy DNA construct." The promoter may be homologous or heterologous to the endogenous ABCL gene, provided that it is active in the cell or tissue type into which the construct will be inserted. Other components of the gene therapy DNA construct  
10 may optionally include DNA molecules designed for site-specific integration (*e.g.*, endogenous sequences useful for homologous recombination), tissue-specific promoters, enhancers or silencers, DNA molecules capable of providing a selective advantage over the parent cell, DNA molecules useful as labels to identify transformed cells, negative selection systems, cell specific binding agents (as, for  
15 example, for cell targeting), cell-specific internalization factors, transcription factors enhancing expression from a vector, and factors enabling vector production.

A gene therapy DNA construct can then be introduced into cells (either *ex vivo* or *in vivo*) using viral or non-viral vectors. One means for introducing the gene therapy DNA construct is by means of viral vectors as described herein. Certain  
20 vectors, such as retroviral vectors, will deliver the DNA construct to the chromosomal DNA of the cells, and the gene can integrate into the chromosomal DNA. Other vectors will function as episomes, and the gene therapy DNA construct will remain in the cytoplasm.

In yet other embodiments, regulatory elements can be included for the  
25 controlled expression of the ABCL gene in the target cell. Such elements are turned on in response to an appropriate effector. In this way, a therapeutic polypeptide can be expressed when desired. One conventional control means involves the use of small molecule dimerizers or rapalogs to dimerize chimeric proteins which contain a small molecule-binding domain and a domain capable of initiating a biological  
30 process, such as a DNA-binding protein or transcriptional activation protein (*see* International Pub. Nos. WO 96/41865, WO 97/31898, and WO 97/31899). The dimerization of the proteins can be used to initiate transcription of the transgene.

An alternative regulation technology uses a method of storing proteins expressed from the gene of interest inside the cell as an aggregate or cluster. The gene of interest is expressed as a fusion protein that includes a conditional aggregation domain that results in the retention of the aggregated protein in the endoplasmic reticulum. The stored proteins are stable and inactive inside the cell. The proteins can be released, however, by administering a drug (*e.g.*, small molecule ligand) that removes the conditional aggregation domain and thereby specifically breaks apart the aggregates or clusters so that the proteins may be secreted from the cell. See Aridor *et al.*, 2000, *Science* 287:816-17 and Rivera *et al.*, 2000, *Science* 287:826-30.

Other suitable control means or gene switches include, but are not limited to, the systems described herein. Mifepristone (RU486) is used as a progesterone antagonist. The binding of a modified progesterone receptor ligand-binding domain to the progesterone antagonist activates transcription by forming a dimer of two transcription factors that then pass into the nucleus to bind DNA. The ligand-binding domain is modified to eliminate the ability of the receptor to bind to the natural ligand. The modified steroid hormone receptor system is further described in U.S. Patent No. 5,364,791 and International Pub. Nos. WO 96/40911 and WO 97/10337.

Yet another control system uses ecdysone (a fruit fly steroid hormone), which binds to and activates an ecdysone receptor (cytoplasmic receptor). The receptor then translocates to the nucleus to bind a specific DNA response element (promoter from ecdysone-responsive gene). The ecdysone receptor includes a transactivation domain, DNA-binding domain, and ligand-binding domain to initiate transcription. The ecdysone system is further described in U.S. Patent No. 5,514,578 and International Pub. Nos. WO 97/38117, WO 96/37609, and WO 93/03162.

Another control means uses a positive tetracycline-controllable transactivator. This system involves a mutated tet repressor protein DNA-binding domain (mutated tet R-4 amino acid changes which resulted in a reverse tetracycline-regulated transactivator protein, *i.e.*, it binds to a tet operator in the presence of tetracycline) linked to a polypeptide which activates transcription. Such systems are described in U.S. Patent Nos. 5,464,758, 5,650,298, and 5,654,168.

Additional expression control systems and nucleic acid constructs are

described in U.S. Patent Nos. 5,741,679 and 5,834,186, to Innovir Laboratories Inc.

*In vivo* gene therapy may be accomplished by introducing the gene encoding ABCL polypeptide into cells via local injection of an ABCL nucleic acid molecule or by other appropriate viral or non-viral delivery vectors. Hefti 1994, *Neurobiology* 5 25:1418-35. For example, a nucleic acid molecule encoding an ABCL polypeptide may be contained in an adeno-associated virus (AAV) vector for delivery to the targeted cells (*see, e.g.*, Johnson, International Pub. No. WO 95/34670; International App. No. PCT/US95/07178). The recombinant AAV genome typically contains AAV inverted terminal repeats flanking a DNA sequence encoding an ABCL 10 polypeptide operably linked to functional promoter and polyadenylation sequences.

Alternative suitable viral vectors include, but are not limited to, retrovirus, adenovirus, herpes simplex virus, lentivirus, hepatitis virus, parvovirus, papovavirus, poxvirus, alphavirus, coronavirus, rhabdovirus, paramyxovirus, and papilloma virus vectors. U.S. Patent No. 5,672,344 describes an *in vivo* viral-mediated gene transfer 15 system involving a recombinant neurotrophic HSV-1 vector. U.S. Patent No. 5,399,346 provides examples of a process for providing a patient with a therapeutic protein by the delivery of human cells that have been treated *in vitro* to insert a DNA segment encoding a therapeutic protein. Additional methods and materials for the practice of gene therapy techniques are described in U.S. Patent Nos. 5,631,236 20 (involving adenoviral vectors), 5,672,510 (involving retroviral vectors), 5,635,399 (involving retroviral vectors expressing cytokines).

Nonviral delivery methods include, but are not limited to, liposome-mediated transfer, naked DNA delivery (direct injection), receptor-mediated transfer (ligand-DNA complex), electroporation, calcium phosphate precipitation, and microparticle 25 bombardment (*e.g.*, gene gun). Gene therapy materials and methods may also include inducible promoters, tissue-specific enhancer-promoters, DNA sequences designed for site-specific integration, DNA sequences capable of providing a selective advantage over the parent cell, labels to identify transformed cells, negative selection systems and expression control systems (safety measures), cell-specific binding 30 agents (for cell targeting), cell-specific internalization factors, and transcription factors to enhance expression by a vector as well as methods of vector manufacture. Such additional methods and materials for the practice of gene therapy techniques are

described in U.S. Patent Nos. 4,970,154 (involving electroporation techniques), 5,679,559 (describing a lipoprotein-containing system for gene delivery), 5,676,954 (involving liposome carriers), 5,593,875 (describing methods for calcium phosphate transfection), and 4,945,050 (describing a process wherein biologically active  
5 particles are propelled at cells at a speed whereby the particles penetrate the surface of the cells and become incorporated into the interior of the cells), and International Pub. No. WO 96/40958 (involving nuclear ligands).

It is also contemplated that ABCL gene therapy or cell therapy can further include the delivery of one or more additional polypeptide(s) in the same or a  
10 different cell(s). Such cells may be separately introduced into the patient, or the cells may be contained in a single implantable device, such as the encapsulating membrane described above, or the cells may be separately modified by means of viral vectors.

A means to increase endogenous ABCL polypeptide expression in a cell via gene therapy is to insert one or more enhancer elements into the ABCL polypeptide  
15 promoter, where the enhancer elements can serve to increase transcriptional activity of the ABCL gene. The enhancer elements used will be selected based on the tissue in which one desires to activate the gene – enhancer elements known to confer promoter activation in that tissue will be selected. For example, if a gene encoding an ABCL polypeptide is to be “turned on” in T-cells, the *lck* promoter enhancer  
20 element may be used. Here, the functional portion of the transcriptional element to be added may be inserted into a fragment of DNA containing the ABCL polypeptide promoter (and optionally, inserted into a vector and/or 5’ and/or 3’ flanking sequences) using standard cloning techniques. This construct, known as a “homologous recombination construct,” can then be introduced into the desired cells  
25 either *ex vivo* or *in vivo*.

Gene therapy also can be used to decrease ABCL polypeptide expression by modifying the nucleotide sequence of the endogenous promoter. Such modification is typically accomplished via homologous recombination methods. For example, a DNA molecule containing all or a portion of the promoter of the ABCL gene selected  
30 for inactivation can be engineered to remove and/or replace pieces of the promoter that regulate transcription. For example, the TATA box and/or the binding site of a transcriptional activator of the promoter may be deleted using standard molecular

biology techniques; such deletion can inhibit promoter activity thereby repressing the transcription of the corresponding ABCL gene. The deletion of the TATA box or the transcription activator binding site in the promoter may be accomplished by generating a DNA construct comprising all or the relevant portion of the ABCL polypeptide promoter (from the same or a related species as the ABCL gene to be regulated) in which one or more of the TATA box and/or transcriptional activator binding site nucleotides are mutated via substitution, deletion and/or insertion of one or more nucleotides. As a result, the TATA box and/or activator binding site has decreased activity or is rendered completely inactive. This construct, which also will typically contain at least about 500 bases of DNA that correspond to the native (endogenous) 5' and 3' DNA sequences adjacent to the promoter segment that has been modified, may be introduced into the appropriate cells (either *ex vivo* or *in vivo*) either directly or via a viral vector as described herein. Typically, the integration of the construct into the genomic DNA of the cells will be via homologous recombination, where the 5' and 3' DNA sequences in the promoter construct can serve to help integrate the modified promoter region via hybridization to the endogenous chromosomal DNA.

#### Therapeutic Uses

ABCL nucleic acid molecules, polypeptides, and agonists and antagonists thereof can be used to treat, diagnose, ameliorate, or prevent a number of diseases, disorders, or conditions, including those recited herein.

ABCL polypeptide agonists and antagonists include those molecules which regulate ABCL polypeptide activity and either increase or decrease at least one activity of the mature form of the ABCL polypeptide. Agonists or antagonists may be co-factors, such as a protein, peptide, carbohydrate, lipid, or small molecular weight molecule, which interact with ABCL polypeptide and thereby regulate its activity. Potential polypeptide agonists or antagonists include antibodies that react with either soluble or membrane-bound forms of ABCL polypeptides that comprise part or all of the extracellular domains of the said proteins. Molecules that regulate ABCL polypeptide expression typically include nucleic acids encoding ABCL polypeptide that can act as anti-sense regulators of expression.

Since ABCL polypeptide shares structural homology with other members of the ABC transporter superfamily (*see* Example 1), it is likely that ABCL polypeptides play a role in the ATP-dependent translocation of solutes across biological membranes. The pattern of expression of ABCL polypeptide (*see* Example 3) also  
5 suggests a role for ABCL polypeptides in various disease states.

The structural similarity of ABCL polypeptide to ABCA1 polypeptide suggests that ABCL polypeptides may be involved in the transport of lipids. Accordingly, ABCL nucleic acid molecules, polypeptides, agonists and antagonists thereof may play a role in the diagnosis and/or treatment of diseases and conditions  
10 involving impaired transport of lipids. Examples of such diseases include, but are not limited to, cardiovascular disease, hypertriglyceridemia, atherosclerosis, hypercholesterolemia, hypocholesterolemia, HDL deficiency syndrome, Tangier disease, and other dyslipidemias and dyslipidemia-related conditions. Other diseases and conditions involving impaired transport of lipids are encompassed by the scope of  
15 this invention. ABCL polypeptides may also be useful in the diagnosis and treatment of cholesterol mediated neurotoxicity.

Due to the structural similarity between ABCL polypeptide and other members of the ABC transporter superfamily, ABCL nucleic acid molecules, polypeptides, agonists and antagonists thereof may also play a role in the diagnosis  
20 and/or treatment of diseases and conditions involving the transport of molecules other than lipids. For example, ABCL polypeptides may be involved in the transport of neurosteroids such as DHEA and progesterone. Thus, ABCL nucleic acid molecules, polypeptides, agonists and antagonists may play a role in the diagnosis and/or treatment of diseases and conditions involving functional and trophic disturbances of  
25 the nervous system. Other examples of diseases that may involve ABCL polypeptide expression include, but are not limited to, Stargardt disease, degenerative and inflammatory retinopathy, cystis fibrosis, and conditions involving multidrug resistance. Other diseases and conditions involving the transport of molecules other than lipids are encompassed by the scope of this invention.

30 The expression of ABCL polypeptide in the thymus and spleen suggests that ABCL polypeptides may play a role in the development and function of these tissues and cellular processes related to these tissues. Accordingly, ABCL nucleic acid

molecules, polypeptides, agonists and antagonists thereof may play a role in the diagnosis and/or treatment of diseases and conditions involving lymphoid and myeloid cells. Examples of such diseases include, but are not limited to, the modulation of immune responses, AIDS, lymphomas, leukemias, neutropenia, anemia, and autoimmune diseases. Other diseases involving hematopoietic cells are encompassed within the scope of the invention.

The expression of ABCL polypeptide in the thyroid suggests that ABCL polypeptides may play a role in the development and function of this tissue. Accordingly, ABCL nucleic acid molecules, polypeptides, agonists and antagonists thereof may play a role in the diagnosis and/or treatment of diseases and conditions involving the thyroid. Examples of such diseases include, but are not limited to, hypothyroidism and hyperthyroidism. Other diseases involving the thyroid are encompassed within the scope of the invention.

The expression of ABCL polypeptide in the hypothalamus suggests that ABCL polypeptides may play a role in the development and function of this tissue. Accordingly, ABCL nucleic acid molecules, polypeptides, agonists and antagonists thereof may play a role in the diagnosis and/or treatment of diseases and conditions involving the hypothalamus. Examples of such diseases include, but are not limited to, obesity, diabetes, reproductive disorders, and energy balance disorders. Other diseases involving the hypothalamus are encompassed within the scope of the invention.

The expression of ABCL polypeptide in the sensory, sympathetic, and parasympathetic ganglia suggests that ABCL polypeptides may play a role in the development and function of these tissues. Accordingly, ABCL nucleic acid molecules, polypeptides, agonists and antagonists thereof may play a role in the diagnosis and/or treatment of peripheral neuropathies including myelinopathies and axonopathies. Examples of such diseases include, but are not limited to, Charcot-Marie-Tooth disease, Dejerine-Sottas syndrome, Guillain-Barr syndrome, and diabetic neuropathy. Also included are the diagnosis and treatment of autoimmune and inflammatory diseases involving the nervous system, including, but not limited to, multiple sclerosis.

Agonists or antagonists of ABCL polypeptide function may be used (simultaneously or sequentially) in combination with one or more cytokines, growth factors, antibiotics, anti-inflammatories, and/or chemotherapeutic agents as is appropriate for the condition being treated.

5 Other diseases or disorders caused by or mediated by undesirable levels of ABCL polypeptides are encompassed within the scope of the invention. Undesirable levels include excessive levels of ABCL polypeptides and sub-normal levels of ABCL polypeptides.

#### 10 Uses of ABCL Nucleic Acids and Polypeptides

Nucleic acid molecules of the invention (including those that do not themselves encode biologically active polypeptides) may be used to map the locations of the ABCL gene and related genes on chromosomes. Mapping may be done by techniques known in the art, such as PCR amplification and *in situ* hybridization.

15 ABCL nucleic acid molecules (including those that do not themselves encode biologically active polypeptides), may be useful as hybridization probes in diagnostic assays to test, either qualitatively or quantitatively, for the presence of an ABCL nucleic acid molecule in mammalian tissue or bodily fluid samples.

20 Other methods may also be employed where it is desirable to inhibit the activity of one or more ABCL polypeptides. Such inhibition may be effected by nucleic acid molecules that are complementary to and hybridize to expression control sequences (triple helix formation) or to ABCL mRNA. For example, antisense DNA or RNA molecules, which have a sequence that is complementary to at least a portion of an ABCL gene can be introduced into the cell. Anti-sense probes may be designed by available techniques using the sequence of the ABCL gene disclosed herein. Typically, each such antisense molecule will be complementary to the start site (5' end) of each selected ABCL gene. When the antisense molecule then hybridizes to the corresponding ABCL mRNA, translation of this mRNA is prevented or reduced.

25 Anti-sense inhibitors provide information relating to the decrease or absence of an ABCL polypeptide in a cell or organism.

Alternatively, gene therapy may be employed to create a dominant-negative

inhibitor of one or more ABCL polypeptides. In this situation, the DNA encoding a mutant polypeptide of each selected ABCL polypeptide can be prepared and introduced into the cells of a patient using either viral or non-viral methods as described herein. Each such mutant is typically designed to compete with  
5 endogenous polypeptide in its biological role.

In addition, an ABCL polypeptide, whether biologically active or not, may be used as an immunogen, that is, the polypeptide contains at least one epitope to which antibodies may be raised. Selective binding agents that bind to an ABCL polypeptide (as described herein) may be used for *in vivo* and *in vitro* diagnostic purposes,  
10 including, but not limited to, use in labeled form to detect the presence of ABCL polypeptide in a body fluid or cell sample. The antibodies may also be used to prevent, treat, or diagnose a number of diseases and disorders, including those recited herein. The antibodies may bind to an ABCL polypeptide so as to diminish or block at least one activity characteristic of an ABCL polypeptide, or may bind to a  
15 polypeptide to increase at least one activity characteristic of an ABCL polypeptide (including by increasing the pharmacokinetics of the ABCL polypeptide).

ABCL polypeptides can be used to clone ABCL ligands using an "expression cloning" strategy. Radiolabeled (<sup>125</sup>Iodine) ABCL polypeptide or "affinity/activity-tagged" ABCL polypeptide (such as an Fc fusion or an alkaline phosphatase fusion)  
20 can be used in binding assays to identify a cell type, cell line, or tissue that expresses an ABCL ligand. RNA isolated from such cells or tissues can then be converted to cDNA, cloned into a mammalian expression vector, and transfected into mammalian cells (*e.g.*, COS or 293) to create an expression library. Radiolabeled or tagged ABCL polypeptide can then be used as an affinity reagent to identify and isolate the  
25 subset of cells in this library expressing an ABCL ligand. DNA is then isolated from these cells and transfected into mammalian cells to create a secondary expression library in which the fraction of cells expressing the ABCL ligand would be many-fold higher than in the original library. This enrichment process can be repeated iteratively until a single recombinant clone containing the ABCL ligand is isolated.  
30 Isolation of ABCL ligands is useful for identifying or developing novel agonists and antagonists of the ABCL signaling pathway. Such agonists and antagonists include ABCL ligands, anti-ABCL ligand antibodies, small molecules or antisense

oligonucleotides.

The human ABCL nucleic acids of the present invention are also useful tools for isolating the corresponding chromosomal ABCL polypeptide genes. The human ABCL genomic DNA can be used to identify heritable tissue-degenerating diseases.

5 Deposits of cDNA encoding murine and human ABCL polypeptide and human ABCL1550, having Accession Nos. PTA-3109, PTA-3110, or PTA-3111, were made with the American Type Culture Collection, 10801 University Boulevard, Manassas, VA 20110-2209 on February 23, 2001.

The following examples are intended for illustration purposes only, and  
10 should not be construed as limiting the scope of the invention in any way.

#### Example 1: Cloning of the Human ABCL Polypeptide Gene

Generally, materials and methods as described in Sambrook *et al. supra* were used to clone and analyze the gene encoding human ABCL polypeptide.

15 To isolate cDNA sequences encoding human ABCL polypeptide, a search of the Celera ORF (Celera, Rockville, MD) and Amgenesis (Amgen, Thousand Oaks, CA) databases was performed using human ABCA1 and ABCR as query sequences. A number of clones from the Celera ORF database were found to contain a portion of the full-length nucleic acid sequence for a novel member of the ABC1 superfamily,  
20 ABCL. A clone from the Amgenesis database (zhgb-al137481) was also found to contain a portion of the full-length ABCL cDNA sequence. The full-length ABCL cDNA sequence was assembled from three partial cDNA clones (RDS# 200020026, RDS# 200015777, and RDS# 200016628).

Clone RDS# 200020026 was generated as follows. Sub-pools of a human  
25 whole brain cDNA library (Edge BioSystems, Gaithersburg, MD) were screened by PCR using amplimers corresponding to the first ATP binding cassette of the human ABCL gene. PCR was performed using the amplimers 2476-84 (5'-T-C-C-A-T-C-T-T-G-A-G-T-G-G-C-C-T-C-T-T-C-C-C-3'; SEQ ID NO: 15) and 2476-85 (5'-C-A-G-A-C-C-C-T-T-C-A-G-C-C-G-C-C-C-A-T-A-G-3'; SEQ ID NO: 16), and a  
30 Boehringer Mannheim Taq polymerase kit. Reactions were performed at 94°C for 2 minutes for one cycle; 94°C for 20 seconds and 72°C for 1 minute for 35 cycles; and 72°C for 5 minutes for one cycle on a Perkin-Elmer (Sunnyvale, CA) 9600 or 2400

thermal cycler. Following amplification, reactions were analyzed on a 1% agarose gel. Sub-pools of the human whole brain cDNA library yielding PCR products of approximately 200 bp were selected for a secondary screening by PCR.

In the secondary screen, selected sub-pools of the cDNA library were  
5 analyzed as described above using the amplimers 2521-12 (5'-A-T-G-G-C-C-T-T-C-T-G-G-A-C-A-C-A-G-C-T-G-A-T-G-3'; SEQ ID NO: 17) and 2508-66 (5'-C-T-T-C-A-G-G-C-G-T-C-T-C-C-A-G-A-G-C-A-G-G-3'; SEQ ID NO: 18). Following amplification, reactions were analyzed on a 1% agarose gel. One sub-pool (2E8), which yielded a PCR product of approximately 1 kb in the secondary screen, was  
10 selected for further amplification experiments.

Clone RDS# 200020026 was isolated from this sub-pool using an ABCL probe (5902) corresponding to the first ATP-binding domain of the human ABCL gene. This probe was generated by PCR from a human stomach Marathon™ cDNA library (Clontech, Palo Alto, CA) using the amplimers 2476-84 and 2476-85 and an  
15 Advantage2™ PCR Cloning kit (Clontech). Reactions were performed at 94°C for 1 minute for one cycle; 94°C for 30 seconds and 72°C for 2 minutes for 5 cycles; 94°C for 30 seconds and 70°C for 2 minutes for 5 cycles; and 94°C for 30 seconds and 68°C for 2 minutes for 25 cycles on a Perkin-Elmer 2400 thermal cycler. Following amplification, reactions were analyzed on a 1% agarose gel. A PCR product of  
20 approximately 200 bp was excised from the agarose gel, purified using a Gel Extraction kit (Qiagen, Chatsworth, CA), and then ligated into pCR2.1 (Invitrogen, Carlsbad, CA). DH5α cells were transformed with the ligation reaction and grown overnight. Plasmid DNA was isolated from the bacterial host cells using a Qiagen miniprep protocol and then analyzed by digestion with Eco RI. A clone containing  
25 an insert of approximately 200 bp was sequenced to verify the presence of the human ABCL cDNA sequence. The probe was isolated by digestion of this clone with Eco RI and purification of the fragment using the Gel Extraction kit. The probe was labeled with α-<sup>32</sup>P-dCTP (RediVue, Amersham, Arlington Heights, IL) using a RediPrime random primed reaction kit (Amersham), and then purified by size  
30 exclusion chromatography (5Prime-3Prime, Boulder, CO).

The 5902 probe was used to screen DH10B electrocompetent cells which were transformed with 0.5 μl of sub-pool 2E8 and grown overnight on LB plates

containing ampicillin. Colonies were transferred to nylon filters, and the filters hybridized for 2 hours at 65°C in 60 ml of RapidHyb buffer (Amersham) containing the labeled probe. Following hybridization, the filters were washed twice for 40 minutes at room temperature in 2X SSC and 0.05% SDS, and then twice for 40 minutes at 65°C in 0.1X SSC and 0.1% SDS. Filters were then exposed to X-ray film for 4 hours. Plasmid DNA for each of eight positive colonies was isolated using the Qiagen miniprep protocol and then digested with Not I and Eco RI. Three colonies were found to contain an insert of approximately 7 kb. One of these clones (RDS# 200020026) was sequenced and found to contain ABCL nucleic acid sequence. However, analysis of this sequence revealed an insertion of approximately 350 bp in the first transmembrane domain, which introduces two stop codons, and a deletion of approximately 270 bp in the second transmembrane domain, which produces a truncated ABCL protein of 1550 amino acids (human ABCL1550).

Clone RDS# 200015777 was generated by PCR using the human stomach Marathon™ cDNA library, the amplimers 2508-63 (5'-G-T-G-C-C-C-T-G-G-C-T-C-C-A-G-G-G-T-C-T-C-3'; SEQ ID NO: 19) and 2477-29 (5'-C-A-G-C-A-C-G-T-A-G-G-G-C-A-G-G-T-A-G-A-G-G-3'; SEQ ID NO: 20), and the Advantage2™ PCR Cloning kit. Reactions were performed at 94°C for 1 minute for one cycle; 94°C for 30 seconds and 72°C for 4 minutes for 5 cycles; 94°C for 30 seconds and 70°C for 4 minutes for 5 cycles; and 94°C for 30 seconds and 68°C for 4 minutes for 5 cycles on a Perkin-Elmer 2400 thermal cycler. Following amplification, reactions were analyzed on a 1% agarose gel. A PCR product of approximately 2 kb was excised from the gel, purified using the Qiagen Gel Extraction kit, and then ligated into pCR2.1. DH5α cells were transformed with the ligation reaction and grown overnight. Plasmid DNA was isolated from the bacterial host cells using the Qiagen miniprep protocol and then analyzed by digestion with Eco RI. A clone (RDS# 200020026) containing an insert of approximately 2 kb was sequenced to verify the presence of the human ABCL cDNA sequence. The ABCL nucleic acid sequence contained within this clone was found to span the first transmembrane domain and lack the insertion identified in clone RDS# 200020026.

Clone RDS# 200016628 was generated by PCR using the human Marathon™ cDNA library (Clontech), the amplimers 2476-84 and 2520-24 (5'-G-A-A-C-C-G-C-

C-C-A-T-T-C-A-C-C-A-T-G-A-T-G-G-3'; SEQ ID NO: 21, derived from clone zhgb-al137481), and the Advantage2™ PCR Cloning kit. Reactions were performed at 94°C for 1 minute for one cycle; 94°C for 30 seconds and 72°C for 5 minutes for 5 cycles; 94°C for 30 seconds and 70°C for 5 minutes for 5 cycles; and 94°C for 30 seconds and 68°C for 5 minutes for 5 cycles on a Perkin-Elmer 2400 thermal cycler. Following amplification, reactions were analyzed on a 1% agarose gel. A PCR product of approximately 3.5 kb was excised from the gel, purified using the Qiagen Gel Extraction kit, and then ligated into pCR2.1. DH5α cells were transformed with the ligation reaction and grown overnight. Plasmid DNA was isolated from the bacterial host cells using the Qiagen miniprep protocol and then analyzed by digestion with Eco RI. A clone (RDS# 200016628) containing an insert of approximately 3.5 kb was sequenced to verify the presence of the human ABCL cDNA sequence. The ABCL nucleic acid sequence contained within this clone was found to span the second transmembrane domain and lack the deletion identified in clone RDS# 200020026.

The full-length sequence of the human ABCL gene sequence was obtained by replacing a Nhe I - Cla I fragment excised from the ABCL cDNA sequence of clone RDS# 200020026 with the Nhe I - Cla I fragment isolated from the ABCL cDNA sequence of clone RDS# 200015777 (thus removing the inserted sequence contained within the first transmembrane domain of clone RDS# 200020026), and then replacing a Hind II - Hind II fragment excised from the ABCL cDNA sequence of clone RDS# 200020026 with the Hind II - Hind II fragment isolated from the ABCL cDNA sequence of clone RDS# 2000216628 (thus replacing the sequence deleted from the second transmembrane domain of clone RDS# 200020026). Although an additional sequence discrepancy was detected between the clones at position 2630 (where the base may be either A or C), this discrepancy does not alter the amino acid sequence of the resulting protein.

To generate a full-length human ABCL cDNA clone, the ABCL cDNA insert from clone RDS# 200020026 was transferred to pCR2.1 as follows. Clone RDS# 200020026 was digested with Eco RI, Not I, and Sal I and electrophoresed on a 1% agarose gel. A single band of approximately 7 kb was excised from the gel, purified using the Qiagen Gel Extraction kit, and then ligated into pCR2.1 (digested with Not

I and Eco RI). DH5 $\alpha$  cells were transformed with the ligation reaction and grown overnight. Plasmid DNA was isolated from the bacterial host cells using the Qiagen miniprep protocol and then analyzed by digestion with Nhe I and Cla I. Plasmid DNA digested with Nhe I and Cla I was then electrophoresed on a 1% agarose gel and a single band of approximately 10 kb was excised from the gel and purified using the Qiagen Gel Extraction kit.

Clone RDS# 200015777 was digested with Nhe I and Cla I and electrophoresed on a 1% agarose gel. A single band of approximately 1 kb was excised from the gel, purified using the Qiagen Gel Extraction kit, and then ligated into the plasmid described above (following digestion with Nhe I and Cla I). DH5 $\alpha$  cells were transformed with the ligation reaction and grown overnight. Plasmid DNA was isolated from the bacterial host cells using the Qiagen miniprep protocol and then analyzed by digestion with Hind II. Plasmid DNA digested with Hind II was then electrophoresed on a 1% agarose gel and a single band of approximately 10 kb was excised from the gel and purified using the Qiagen Gel Extraction kit.

Clone RDS# 200016628 was digested with Hind II and electrophoresed on a 1% agarose gel. A single band of approximately 1 kb was excised from the gel, purified using the Qiagen Gel Extraction kit, and then ligated into the plasmid described above (following digestion with Hind II). DH5 $\alpha$  cells were transformed with the ligation reaction and grown overnight. Plasmid DNA containing the full-length sequence of the human ABCL gene (pCR2.1-huABCL) was isolated from the bacterial host cells using the Qiagen miniprep protocol.

Sequence analysis of the predicted cDNA sequence for human ABCL polypeptide indicated that the gene comprises a 6438 bp open reading frame encoding a protein of 2146 amino acids (Figures 2A-2K; predicted signal peptide indicated by underline). Sequence analysis indicated that ABCL polypeptide was more specifically related to the ABCA subfamily of the ABC Transporter superfamily. Figures 5A-5E illustrate an amino acid sequence alignment of human ABC1 polypeptide (huABC1; GenBank Accession No. AAF86276; SEQ ID NO: 11), human ABCR polypeptide (huABCR; GenBank Accession No. CAA75729; SEQ ID NO: 12), and human ABCL polypeptide (huABCL; SEQ ID NO: 5).

Sequence analysis of the predicted cDNA sequence for the truncated form of the human ABCL polypeptide (ABCL1550) indicated that the gene comprises a 4650 bp open reading frame encoding a protein of 1550 amino acids (Figures 3A-3H).

5 ABCL is similar to ABC transporters of the ABC1 subfamily, which includes ABC1, ABC2, ABC-C, and ABCR. ABCL is most closely related to ABC1, with which it shares 54% amino acid sequence identity, and ABCR, with which it shares 49% amino acid sequence identity.

Broccardo *et al.* (*Biochim. Biophys. Acta* 1461:395-404 (1999)) teach an ABC  
10 transporter protein that is preferentially expressed in the spleen and more generally expressed in other lymphoid organs. The authors also teach the chromosomal location (10B4-C1) of the gene encoding this member of the ABC transporter superfamily. Finally, Broccardo *et al.* teach that the gene encoding this member of the ABC transporter superfamily comprises more than 37 exons. However, the  
15 authors do not teach the nucleic acid sequence encoding this polypeptide. While Broccardo *et al.* refer to this member of the ABC transporter superfamily as ABCA7, ABC4, and ABCX, the gene that the authors teach shares some similarity with ABCL polypeptide.

Schmitz and Klucken (PCT Publication No. WO 00/18912) teach an mRNA  
20 transcript of approximately 6.5 kb encoding an ABC transporter protein. The authors also teach that the gene encoding this member of the ABC transporter superfamily maps to Ch 19p13. Schmitz and Klucken further teach the up- regulation of this member of the ABC transporter superfamily by cholesterol loading and the down-regulation of the protein by cholesterol depletion. Finally, while Schmitz and  
25 Klucken teach a cDNA fragment of 2911 bp encoding a portion of this ABC transporter protein (which the authors refer to as ABCA8 or ABC-new), Schmitz and Klucken do not teach the entire nucleic acid sequence encoding this polypeptide. The cDNA transcript taught by the authors encodes approximately half of this ABC transporter protein (the partial sequence of which shares sequence identity with the  
30 ABCL gene disclosed herein). While the investigators provide similar teachings in Klucken *et al.*, 2000, *Proc. Natl. Acad. Sci. U.S.A.* 97:817-22, the authors refer to the ABC transporter protein as ABCA7 and ABC4.

Schriml and Dean (*Genomics* 64:24-31 (2000)) teach an 863 bp expressed sequence tag (EST) sequence (EST665146) encoding a portion of the ABCA7 polypeptide. The authors also teach that ABCA7 (which the authors also refer to as Abc51) is expressed in monocytes and macrophages. Similarly to Schmitz and Klucken, Schriml and Dean teach the up-regulation of ABCA7 by cholesterol loading and the down-regulation of ABCA7 by cholesterol depletion. The authors further teach that formation of cholesterol gall stones in mice can be linked to the *Lith1* region of chromosome 10 and that the ABCA7 gene maps to this region of the chromosome (Ch 10 44cM).

10 Kaminski *et al.* (Biochem. Biophys. Res. Commun. 273:532-38 (2000)) teach the nucleic acid sequence for an ABC transporter protein they designate as ABCA7. The authors also teach an mRNA transcript of approximately 6.8 kb encoding this protein. Kaminski *et al.* further teach that ABCA7 polypeptide shares the highest sequence identity with the human cholesterol and phospholipid exporter ABCA1  
15 (54%) and the human retinal transporter ABCR (49%). Finally, the authors teach that ABCA7 mRNA was detected predominantly in myelo-lymphatic tissues, with the highest expression in peripheral leukocytes, thymus, spleen, and bone marrow.

#### Example 2: Cloning of the Murine ABCL Polypeptide Gene

20 Generally, materials and methods as described in Sambrook *et al. supra* were used to clone and analyze the gene encoding murine ABCL polypeptide.

To isolate cDNA sequences encoding murine ABCL polypeptide, a search of the Amgen database was performed using murine ABCA1 and human ABCL cDNA query sequences. Clone smbo2-90-h5 was found to contain a portion of the murine ABCL cDNA sequence and clone srpg2-36-h3 was found to contain a portion  
25 of the rat ABCL cDNA sequence. Clone smbo2-90-h5 was isolated from a directional oligo-dT primed cDNA library prepared from mouse bone (hip to ankle) from HECR1 OPG+ mice using the SuperScript™ cDNA Plasmid System (Gibco-BRL). Sequence analysis indicated that the clone contained an insert of 5 kb  
30 encoding a portion of the murine ABCL gene (spanning from the latter half of the first transmembrane domain to the poly-A tail). Clone srpg2-36-h3 was isolated from a directional random primed cDNA library prepared from rat pineal gland using the

SuperScript™ cDNA Plasmid System. Sequence analysis indicated that the clone contained an insert of 2 kb encoding a portion of the rat ABCL gene (from the 5' end of the gene).

The full-length murine ABCL cDNA sequence was assembled from four  
5 partial cDNA clones (RDS#200019182, RDS#200019183, RDS#200019184, and  
RDS#200019185). Clones RDS#200019182 and RDS#200019183 were generated  
by PCR using a mouse E17 Marathon™ cDNA library (Clontech), the amplimers  
2533-69 (5'-A-T-G-G-C-T-T-T-C-T-G-C-A-C-A-C-A-G-T-T-G-A-T-G-C-T-C-3';  
SEQ ID NO: 22; derived from clone srpg2-36-h3) and 2507-27 (5'-C-C-A-G-G-C-C-  
10 A-C-A-C-A-C-A-G-T-A-C-G-T-A-G-G-3'; SEQ ID NO: 23; derived from clone  
smbo2-90-h5), and the Advantage2™ PCR Cloning kit. Reactions were performed at  
94°C for 1 minute for one cycle; 94°C for 30 seconds and 72°C for 5 minutes for 5  
cycles; 94°C for 30 seconds and 70°C for 5 minutes for 5 cycles; and 94°C for 30  
seconds and 68°C for 5 minutes for 25 cycles on a Perkin-Elmer 2400 thermal cycler.  
15 Following amplification, reactions were analyzed on a 1% agarose gel. However, no  
PCR products were detected.

On re-amplification of the mixture using the amplimers 2533-69 and 2507-28  
(5'-A-C-A-C-A-G-T-A-C-G-T-A-G-G-G-C-A-G-A-T-A-G-A-G-3'; SEQ ID NO: 24;  
derived from clone smbo2-90-h5), a single band migrating at approximately 3.5 kb  
20 was obtained. This band was separated by gel eletrophoresis, excised from the gel,  
purified using the Qiagen Gel Extraction kit, and then ligated into pCR2.1. DH5α  
cells were transformed with the ligation reaction and grown overnight. Plasmid DNA  
was isolated from the bacterial host cells using the Qiagen miniprep protocol and then  
analyzed by digestion with Eco RI. Two clones containing an insert of approximately  
25 3.5 kb were sequenced and found to contain the 5' end of the murine ABCL cDNA  
sequence.

Clones RDS#200019184 and RDS#200019185 were generated by PCR using  
a mouse lung Marathon™ cDNA library (Clontech), the amplimers 2533-69 and  
2507-27, and the Advantage2™ PCR Cloning kit. Reactions were performed at 94°C  
30 for 1 minute for one cycle; 94°C for 30 seconds and 72°C for 5 minutes for 5 cycles;  
94°C for 30 seconds and 70°C for 5 minutes for 5 cycles; and 94°C for 30 seconds  
and 68°C for 5 minutes for 25 cycles on a Perkin-Elmer 2400 thermal cycler.

Following amplification, reactions were analyzed on a 1% agarose gel. However, no PCR products were detected.

On re-amplification of the mixture using the amplimers 2533-69 and 2507-28, a single band migrating at approximately 3.5 kb was obtained. This band was separated by gel electrophoresis, excised from the gel, purified using the Qiagen Gel Extraction kit, and then ligated into pCR2.1. DH5 $\alpha$  cells were transformed with the ligation reaction and grown overnight. Plasmid DNA was isolated from the bacterial host cells using the Qiagen miniprep protocol and then analyzed by digestion with Eco RI. Two clones containing an insert of approximately 3.5 kb were sequenced and found to contain the 5' end of the murine ABCL cDNA sequence.

A consensus sequence, generated from clones RDS# 200019182, RDS# 200019183, RDS# 200019184, and RDS# 200019185, was used for the 5' end sequence of the murine ABCL gene. Near the end of the first transmembrane domain, an overlap between the PCR generated sequences and that of clone smbo2-90-h5 was observed. Since the sequence of clone smbo2-90-h5 sequence extends from the first transmembrane domain of the murine ABCL gene to the poly-A tail of this gene, the full-length sequence of the murine ABCL was constructed by assembling the fragments generated from PCR amplification with the partial murine ABCL cDNA sequence of clone smbo2-90-h5.

Sequence analysis of the predicted cDNA sequence for murine ABCL polypeptide indicated that the gene comprises a 6501 bp open reading frame encoding a protein of 2167 amino acids (Figures 1A-1K; predicted signal peptide indicated by underline). Figures 4A-4E illustrate an amino acid sequence alignment of murine ABC1 polypeptide (muABC1; GenBank Accession No. NP\_038482; SEQ ID NO: 9), murine ABCR polypeptide (muABCR; GenBank Accession No. NP\_031404; SEQ ID NO: 10), and murine ABCL polypeptide (muABCL; SEQ ID NO: 2).

### Example 3: ABCL mRNA Expression

Multiple human and murine tissue Northern blots (Clontech) were hybridized with a probe generated by PCR amplification of human or murine ABCL cDNA using appropriate amplimers. The probe was radioactively labeled using a Redi-Prime II kit (Amersham) according to the manufacturer's instructions. Northern blots

were prehybridized for 30 minutes at 65°C in Rapid-Hyb buffer (Amersham), and then hybridized in a hybridization oven (Stratagene) for 90 minutes at 65°C in Rapid-Hyb buffer containing the radioactively labeled probe. Following hybridization, the blots were washed twice in 2X SSC and 0.05% SDS at room temperature and then  
5 twice in 0.1X SSC and 0.1% SDS for 1 hour at 65°C. Hybridized blots were examined by phosphorimager.

Northern blot analysis of human tissues indicated that ABCL mRNA was highly expressed in the thymus. Levels of expression above background were also observed in human thyroid and hypothalamus. Northern blot analysis of murine  
10 tissues indicated that ABCL mRNA was expressed in murine spleen, brain and lung. Low levels of expression were also observed in murine with heart and testis.

The expression of ABCL mRNA was localized by *in situ* hybridization. A panel of normal embryonic and adult murine and monkey tissues was fixed in 4%  
15 paraformaldehyde, embedded in paraffin, and sectioned at 5 µm. Sectioned tissues were permeabilized in 0.2 M HCl, digested with Proteinase K, and acetylated with triethanolamine and acetic anhydride. Sections were prehybridized for 1 hour at 60°C in hybridization solution (300 mM NaCl, 20 mM Tris-HCl, pH 8.0, 5 mM EDTA, 1X Denhardt's solution, 0.2% SDS, 10 mM DTT, 0.25 mg/ml tRNA, 25 µg/ml polyA, 25  
20 µg/ml polyC and 50% formamide) and then hybridized overnight at 60°C in the same solution containing 10% dextran and  $2 \times 10^4$  cpm/µl of a <sup>33</sup>P-labeled antisense riboprobe complementary to either the murine ABCL gene (for analysis of murine tissue sections) or the human ABCL gene (for analysis of monkey tissue sections). The riboprobe was obtained by *in vitro* transcription of a clone containing the  
25 appropriate ABCL cDNA sequence using standard techniques.

Following hybridization, sections were rinsed in hybridization solution, treated with RNaseA to digest unhybridized probe, and then washed in 0.1X SSC at 55°C for 30 minutes. Sections were then immersed in NTB-2 emulsion (Kodak, Rochester, NY), exposed for 3 weeks at 4°C, developed, and counterstained with  
30 hematoxylin and eosin. Tissue morphology and hybridization signal were simultaneously analyzed by darkfield and standard illumination for brain (one sagittal and two coronal sections), gastrointestinal tract (esophagus, stomach, duodenum,

jejunum, ileum, proximal colon, and distal colon), pituitary, liver, lung, heart, spleen, thymus, lymph nodes, kidney, adrenal, bladder, pancreas, salivary gland, male and female reproductive organs (ovary, oviduct, and uterus in the female; and testis, epididymus, prostate, seminal vesicle, and vas deferens in the male), BAT and WAT  
5 (subcutaneous, peri-renal), bone (femur), skin, breast, and skeletal muscle.

Expression of ABCL mRNA was detected in both murine and monkey sympathetic and parasympathetic ganglia. In murine embryos, a positive signal was seen in the dorsal root ganglion, superior cervical ganglion, and cochlear ganglion. A signal not attributable to any one specific cell type was also seen in thymus, spleen,  
10 lymph nodes, and lung of both species.

#### Example 4: Production of ABCL Polypeptides

##### A. Expression of ABCL Polypeptides in Bacteria

PCR is used to amplify template DNA sequences encoding an ABCL  
15 polypeptide using primers corresponding to the 5' and 3' ends of the sequence. The amplified DNA products may be modified to contain restriction enzyme sites to allow for insertion into expression vectors. PCR products are gel purified and inserted into expression vectors using standard recombinant DNA methodology. An exemplary vector, such as pAMG21 (ATCC no. 98113) containing the lux promoter and a gene  
20 encoding kanamycin resistance is digested with Bam HI and Nde I for directional cloning of inserted DNA. The ligated mixture is transformed into an *E. coli* host strain by electroporation and transformants are selected for kanamycin resistance. Plasmid DNA from selected colonies is isolated and subjected to DNA sequencing to confirm the presence of the insert.

25 Transformed host cells are incubated in 2xYT medium containing 30 µg/mL kanamycin at 30°C prior to induction. Gene expression is induced by the addition of N-(3-oxohexanoyl)-dl-homoserine lactone to a final concentration of 30 ng/mL followed by incubation at either 30°C or 37°C for six hours. The expression of ABCL polypeptide is evaluated by centrifugation of the culture, resuspension and lysis of the  
30 bacterial pellets, and analysis of host cell proteins by SDS-polyacrylamide gel electrophoresis.

Inclusion bodies containing ABCL polypeptide are purified as follows. Bacterial cells are pelleted by centrifugation and resuspended in water. The cell suspension is lysed by sonication and pelleted by centrifugation at 195,000 xg for 5 to 10 minutes. The supernatant is discarded, and the pellet is washed and transferred to  
5 a homogenizer. The pellet is homogenized in 5 mL of a Percoll solution (75% liquid Percoll and 0.15 M NaCl) until uniformly suspended and then diluted and centrifuged at 21,600 xg for 30 minutes. Gradient fractions containing the inclusion bodies are recovered and pooled. The isolated inclusion bodies are analyzed by SDS-PAGE.

A single band on an SDS polyacrylamide gel corresponding to *E. coli*-  
10 produced ABCL polypeptide is excised from the gel, and the N-terminal amino acid sequence is determined essentially as described by Matsudaira *et al.*, 1987, *J. Biol. Chem.* 262:10-35.

#### B. Expression of ABCL Polypeptide in Mammalian Cells

15 PCR is used to amplify template DNA sequences encoding an ABCL polypeptide using primers corresponding to the 5' and 3' ends of the sequence. The amplified DNA products may be modified to contain restriction enzyme sites to allow for insertion into expression vectors. PCR products are gel purified and inserted into expression vectors using standard recombinant DNA methodology. An exemplary  
20 expression vector, pCEP4 (Invitrogen, Carlsbad, CA), that contains an Epstein-Barr virus origin of replication, may be used for the expression of ABCL polypeptides in 293-EBNA-1 cells. Amplified and gel purified PCR products are ligated into pCEP4 vector and introduced into 293-EBNA cells by lipofection. The transfected cells are selected in 100 µg/mL hygromycin and the resulting drug-resistant cultures are grown  
25 to confluence. The cells are then cultured in serum-free media for 72 hours. The conditioned media is removed and ABCL polypeptide expression is analyzed by SDS-PAGE.

ABCL polypeptide expression may be detected by silver staining. Alternatively, ABCL polypeptide is produced as a fusion protein with an epitope tag,  
30 such as an IgG constant domain or a FLAG epitope, which may be detected by Western blot analysis using antibodies to the peptide tag.

ABCL polypeptides may be excised from an SDS-polyacrylamide gel, or ABCL fusion proteins are purified by affinity chromatography to the epitope tag, and subjected to N-terminal amino acid sequence analysis as described herein.

5 C. Expression and Purification of ABCL Polypeptide in Mammalian Cells

ABCL polypeptide expression constructs are introduced into 293 EBNA or CHO cells using either a lipofection or calcium phosphate protocol.

To conduct functional studies on the ABCL polypeptides that are produced, large quantities of conditioned media are generated from a pool of hygromycin  
10 selected 293 EBNA clones. The cells are cultured in 500 cm Nunc Triple Flasks to 80% confluence before switching to serum free media a week prior to harvesting the media. Conditioned media is harvested and frozen at -20°C until purification.

Conditioned media is purified by affinity chromatography as described below. The media is thawed and then passed through a 0.2 µm filter. A Protein G column is  
15 equilibrated with PBS at pH 7.0, and then loaded with the filtered media. The column is washed with PBS until the absorbance at A<sub>280</sub> reaches a baseline. ABCL polypeptide is eluted from the column with 0.1 M Glycine-HCl at pH 2.7 and immediately neutralized with 1 M Tris-HCl at pH 8.5. Fractions containing ABCL polypeptide are pooled, dialyzed in PBS, and stored at -70°C.

20 For Factor Xa cleavage of the human ABCL polypeptide-Fc fusion polypeptide, affinity chromatography-purified protein is dialyzed in 50 mM Tris-HCl, 100 mM NaCl, 2 mM CaCl<sub>2</sub> at pH 8.0. The restriction protease Factor Xa is added to the dialyzed protein at 1/100 (w/w) and the sample digested overnight at room temperature.

25

Example 5: Production of Anti-ABCL Polypeptide Antibodies

Antibodies to ABCL polypeptides may be obtained by immunization with purified protein or with ABCL peptides produced by biological or chemical synthesis. Suitable procedures for generating antibodies include those described in Hudson and  
30 Bay, *Practical Immunology* (2nd ed., Blackwell Scientific Publications).

In one procedure for the production of antibodies, animals (typically mice or rabbits) are injected with an ABCL antigen (such as an ABCL polypeptide), and those

with sufficient serum titer levels as determined by ELISA are selected for hybridoma production. Spleens of immunized animals are collected and prepared as single cell suspensions from which splenocytes are recovered. The splenocytes are fused to mouse myeloma cells (such as Sp2/0-Ag14 cells), are first incubated in DMEM with  
5 200 U/mL penicillin, 200 µg/mL streptomycin sulfate, and 4 mM glutamine, and are then incubated in HAT selection medium (hypoxanthine, aminopterin, and thymidine). After selection, the tissue culture supernatants are taken from each fusion well and tested for anti-ABCL antibody production by ELISA.

Alternative procedures for obtaining anti-ABCL antibodies may also be  
10 employed, such as the immunization of transgenic mice harboring human Ig loci for production of human antibodies, and the screening of synthetic antibody libraries, such as those generated by mutagenesis of an antibody variable domain.

#### Example 6: Expression of ABCL Polypeptide in Transgenic Mice

15 To assess the biological activity of ABCL polypeptide, a construct encoding an ABCL polypeptide/Fc fusion protein under the control of a liver specific ApoE promoter is prepared. The delivery of this construct is expected to cause pathological changes that are informative as to the function of ABCL polypeptide. Similarly, a construct containing the full-length ABCL polypeptide under the control of the beta  
20 actin promoter is prepared. The delivery of this construct is expected to result in ubiquitous expression.

To generate these constructs, PCR is used to amplify template DNA sequences encoding an ABCL polypeptide using primers that correspond to the 5' and 3' ends of the desired sequence and which incorporate restriction enzyme sites to  
25 permit insertion of the amplified product into an expression vector. Following amplification, PCR products are gel purified, digested with the appropriate restriction enzymes, and ligated into an expression vector using standard recombinant DNA techniques. For example, amplified ABCL polypeptide sequences can be cloned into an expression vector under the control of the human  $\beta$ -actin promoter as described by  
30 Graham *et al.*, 1997, *Nature Genetics*, 17:272-74 and Ray *et al.*, 1991, *Genes Dev.* 5:2265-73.

Following ligation, reaction mixtures are used to transform an *E. coli* host strain by electroporation and transformants are selected for drug resistance. Plasmid DNA from selected colonies is isolated and subjected to DNA sequencing to confirm the presence of an appropriate insert and absence of mutation. The ABCL polypeptide expression vector is purified through two rounds of CsCl density gradient centrifugation, cleaved with a suitable restriction enzyme, and the linearized fragment containing the ABCL polypeptide transgene is purified by gel electrophoresis. The purified fragment is resuspended in 5 mM Tris, pH 7.4, and 0.2 mM EDTA at a concentration of 2 mg/mL.

Single-cell embryos from BDF1 x BDF1 bred mice are injected as described (International Pub. No. WO 97/23614). Embryos are cultured overnight in a CO<sub>2</sub> incubator and 15-20 two-cell embryos are transferred to the oviducts of a pseudopregnant CD1 female mice. Offspring obtained from the implantation of microinjected embryos are screened by PCR amplification of the integrated transgene in genomic DNA samples as follows. Ear pieces are digested in 20 mL ear buffer (20 mM Tris, pH 8.0, 10 mM EDTA, 0.5% SDS, and 500 mg/mL proteinase K) at 55°C overnight. The sample is then diluted with 200 mL of TE, and 2 mL of the ear sample is used in a PCR reaction using appropriate primers.

At 8 weeks of age, transgenic founder animals and control animals are sacrificed for necropsy and pathological analysis. Portions of spleen are removed and total cellular RNA isolated from the spleens using the Total RNA Extraction Kit (Qiagen) and transgene expression determined by RT-PCR. RNA recovered from spleens is converted to cDNA using the SuperScript™ Preamplification System (Gibco-BRL) as follows. A suitable primer, located in the expression vector sequence and 3' to the ABCL polypeptide transgene, is used to prime cDNA synthesis from the transgene transcripts. Ten mg of total spleen RNA from transgenic founders and controls is incubated with 1 mM of primer for 10 minutes at 70°C and placed on ice. The reaction is then supplemented with 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2.5 mM MgCl<sub>2</sub>, 10 mM of each dNTP, 0.1 mM DTT, and 200 U of SuperScript II reverse transcriptase. Following incubation for 50 minutes at 42°C, the reaction is stopped by heating for 15 minutes at 72°C and digested with 2U of RNase

H for 20 minutes at 37°C. Samples are then amplified by PCR using primers specific for ABCL polypeptide.

Example 7: Biological Activity of ABCL Polypeptide in Transgenic Mice

5 Prior to euthanasia, transgenic animals are weighed, anesthetized by isoflurane and blood drawn by cardiac puncture. The samples are subjected to hematology and serum chemistry analysis. Radiography is performed after terminal exsanguination. Upon gross dissection, major visceral organs are subject to weight analysis.

10 Following gross dissection, tissues (*i.e.*, liver, spleen, pancreas, stomach, the entire gastrointestinal tract, kidney, reproductive organs, skin and mammary glands, bone, brain, heart, lung, thymus, trachea, esophagus, thyroid, adrenals, urinary bladder, lymph nodes and skeletal muscle) are removed and fixed in 10% buffered Zn-Formalin for histological examination. After fixation, the tissues are processed  
15 into paraffin blocks, and 3 mm sections are obtained. All sections are stained with hematoxylin and eosin, and are then subjected to histological analysis.

The spleen, lymph node, and Peyer's patches of both the transgenic and the control mice are subjected to immunohistology analysis with B cell and T cell specific antibodies as follows. The formalin fixed paraffin embedded sections are  
20 deparaffinized and hydrated in deionized water. The sections are quenched with 3% hydrogen peroxide, blocked with Protein Block (Lipshaw, Pittsburgh, PA), and incubated in rat monoclonal anti-mouse B220 and CD3 (Harlan, Indianapolis, IN). Antibody binding is detected by biotinylated rabbit anti-rat immunoglobulins and peroxidase conjugated streptavidin (BioGenex, San Ramon, CA) with DAB as a  
25 chromagen (BioTek, Santa Barbara, CA). Sections are counterstained with hematoxylin.

After necropsy, MLN and sections of spleen and thymus from transgenic animals and control littermates are removed. Single cell suspensions are prepared by gently grinding the tissues with the flat end of a syringe against the bottom of a 100  
30 mm nylon cell strainer (Becton Dickinson, Franklin Lakes, NJ). Cells are washed twice, counted, and approximately  $1 \times 10^6$  cells from each tissue are then incubated for 10 minutes with 0.5  $\mu$ g CD16/32(Fc $\gamma$ III/II) Fc block in a 20  $\mu$ L volume. Samples

- are then stained for 30 minutes at 2-8°C in a 100 µL volume of PBS (lacking Ca<sup>+</sup> and Mg<sup>+</sup>), 0.1% bovine serum albumin, and 0.01% sodium azide with 0.5 µg antibody of FITC or PE-conjugated monoclonal antibodies against CD90.2 (Thy-1.2), CD45R (B220), CD11b (Mac-1), Gr-1, CD4, or CD8 (PharMingen, San Diego, CA).
- 5 Following antibody binding, the cells are washed and then analyzed by flow cytometry on a FACScan (Becton Dickinson).

While the present invention has been described in terms of the preferred embodiments, it is understood that variations and modifications will occur to those skilled in the art. Therefore, it is intended that the appended claims cover all such equivalent variations that come within the scope of the invention as claimed.

10

## WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:
  - 5 (a) the nucleotide sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ ID NO: 7;
  - (b) the nucleotide sequence of the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110, or PTA-3111;
  - (c) a nucleotide sequence encoding the polypeptide as set forth in any of  
10 SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;
  - (d) a nucleotide sequence that hybridizes under at least moderately stringent conditions to the complement of the nucleotide sequence of any of (a) - (c); and
  - (e) a nucleotide sequence complementary to the nucleotide sequence of  
15 any of (a) - (c).
2. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:
  - 20 (a) a nucleotide sequence encoding a polypeptide that is at least about 70 percent identical to the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;
  - (b) a nucleotide sequence encoding an allelic variant or splice variant of the nucleotide sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ  
25 ID NO: 7, the nucleotide sequence of the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110, or PTA-3111, or the nucleotide sequence of (a);
  - (c) a region of the nucleotide sequence of any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ ID NO: 7, the nucleotide sequence of the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110, or PTA-3111, or the nucleotide sequence of (a)  
30 or (b) encoding a polypeptide fragment of at least about 25 amino acid residues, wherein the polypeptide fragment has an activity of the encoded polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, or is antigenic;

(d) a region of the nucleotide sequence of any of SEQ ID NO: 1, SEQ ID NO: 4, or SEQ ID NO: 7, the nucleotide sequence of the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110, or PTA-3111, or the nucleotide sequence of any of (a) - (c) comprising a fragment of at least about 16 nucleotides;

5 (e) a nucleotide sequence that hybridizes under at least moderately stringent conditions to the complement of the nucleotide sequence of any of (a) - (d); and

(f) a nucleotide sequence complementary to the nucleotide sequence of any of (a) - (d).

10

3. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one conservative  
15 amino acid substitution, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

(b) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one amino acid  
20 insertion, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

(c) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one amino acid deletion, wherein the encoded polypeptide has an activity of the polypeptide set forth  
in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

25 (d) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 that has a C- and/or N- terminal truncation, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

(e) a nucleotide sequence encoding a polypeptide as set forth in any of  
30 SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions, C-terminal truncation, and N-terminal truncation, wherein the

encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

(f) a nucleotide sequence of any of (a) - (e) comprising a fragment of at least about 16 nucleotides;

5 (g) a nucleotide sequence that hybridizes under at least moderately stringent conditions to the complement of the nucleotide sequence of any of (a) - (f); and

(h) a nucleotide sequence complementary to the nucleotide sequence of any of (a) - (e).

10

4. A vector comprising the nucleic acid molecule of any of Claims 1, 2, or 3.

5. A host cell comprising the vector of Claim 4.

15

6. The host cell of Claim 5 that is a eukaryotic cell.

7. The host cell of Claim 5 that is a prokaryotic cell.

20

8. A process of producing an ABCL polypeptide comprising culturing the host cell of Claim 5 under suitable conditions to express the polypeptide, and optionally isolating the polypeptide from the culture.

9. A polypeptide produced by the process of Claim 8.

25

10. The process of Claim 8, wherein the nucleic acid molecule comprises promoter DNA other than the promoter DNA for the native ABCL polypeptide operatively linked to the DNA encoding the ABCL polypeptide.

30

11. The isolated nucleic acid molecule according to Claim 2, wherein the percent identity is determined using a computer program selected from the group

consisting of GAP, BLASTN, FASTA, BLASTA, BLASTX, BestFit, and the Smith-Waterman algorithm.

12. A process for determining whether a compound inhibits ABCL polypeptide activity or ABCL polypeptide production comprising exposing a cell according to any of Claims 5, 6, or 7 to the compound and measuring ABCL polypeptide activity or ABCL polypeptide production in said cell.

13. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8; and

(b) the amino acid sequence encoded by the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110, or PTA-3111.

15

14. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence as set forth in either SEQ ID NO: 3 or SEQ ID NO: 6, optionally further comprising an amino-terminal methionine;

20 (b) an amino acid sequence for an ortholog of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

(c) an amino acid sequence that is at least about 70 percent identical to the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

25 (d) a fragment of the amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 comprising at least about 25 amino acid residues, wherein the fragment has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, or is antigenic; and

30 (e) an amino acid sequence for an allelic variant or splice variant of the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID

NO: 8, the nucleotide sequence of the DNA insert in ATCC Deposit Nos. PTA-3109, PTA-3110, or PTA-3111, or the amino acid sequence of any of (a) - (c).

15. An isolated polypeptide comprising an amino acid sequence selected  
5 from the group consisting of:

(a) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one conservative amino acid substitution, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

10 (b) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one amino acid insertion, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

(c) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one amino acid deletion, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8;

(d) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 that has a C- and/or N- terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8; and

(e) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8 with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions, C-terminal truncation, and N-terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8.

16. An isolated polypeptide encoded by the nucleic acid molecule of any of Claims 1, 2, or 3, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8.

17. The isolated polypeptide according to Claim 14, wherein the percent identity is determined using a computer program selected from the group consisting of GAP, BLASTP, FASTA, BLASTA, BLASTX, BestFit, and the Smith-Waterman algorithm.

5

18. A selective binding agent or fragment thereof that specifically binds the polypeptide of any of Claims 13, 14, or 15.

19. The selective binding agent or fragment thereof of Claim 18 that specifically binds the polypeptide comprising the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8, or a fragment thereof.

10

20. The selective binding agent of Claim 18 that is an antibody or fragment thereof.

15

21. The selective binding agent of Claim 18 that is a humanized antibody.

22. The selective binding agent of Claim 18 that is a human antibody or fragment thereof.

20

23. The selective binding agent of Claim 18 that is a polyclonal antibody or fragment thereof.

24. The selective binding agent Claim 18 that is a monoclonal antibody or fragment thereof.

25

25. The selective binding agent of Claim 18 that is a chimeric antibody or fragment thereof.

26. The selective binding agent of Claim 18 that is a CDR-grafted antibody or fragment thereof.

30

27. The selective binding agent of Claim 18 that is an antiidiotypic antibody or fragment thereof.
28. The selective binding agent of Claim 18 that is a variable region  
5 fragment.
29. The variable region fragment of Claim 28 that is a Fab or a Fab' fragment.
30. A selective binding agent or fragment thereof comprising at least one  
10 complementarity determining region with specificity for a polypeptide having the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8.
31. The selective binding agent of Claim 18 that is bound to a detectable  
15 label.
32. The selective binding agent of Claim 18 that antagonizes ABCL polypeptide biological activity.
33. A method for treating, preventing, or ameliorating an ABCL  
20 polypeptide-related disease, condition, or disorder comprising administering to a patient an effective amount of a selective binding agent according to Claim 18.
34. A selective binding agent produced by immunizing an animal with a  
25 polypeptide comprising an amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 5, or SEQ ID NO: 8.
35. A hybridoma that produces a selective binding agent capable of  
binding a polypeptide according to any of Claims 13, 14, or 15.  
30
36. A method of detecting or quantitating the amount of ABCL polypeptide using the anti-ABCL antibody or fragment of Claim 18.

37. A kit for detecting or quantitating the amount of ABCL polypeptide in a biological sample, comprising the selective binding agent of Claim 18.

5 38. A composition comprising the polypeptide of any of Claims 13, 14, or 15, and a pharmaceutically acceptable formulation agent.

39. The composition of Claim 38, wherein the pharmaceutically acceptable formulation agent is a carrier, adjuvant, solubilizer, stabilizer, or anti-oxidant.  
10

40. The composition of Claim 39, wherein the polypeptide comprises the amino acid sequence as set forth in either SEQ ID NO: 3 or SEQ ID NO: 6.

15 41. A polypeptide comprising a derivative of the polypeptide of any of Claims 13, 14, or 15.

42. The polypeptide of Claim 41 that is covalently modified with a water-soluble polymer.  
20

43. The polypeptide of Claim 42, wherein the water-soluble polymer is selected from the group consisting of polyethylene glycol, monomethoxy-polyethylene glycol, dextran, cellulose, poly-(N-vinyl pyrrolidone) polyethylene glycol, propylene glycol homopolymers, polypropylene oxide/ethylene oxide co-polymers, polyoxyethylated polyols, and polyvinyl alcohol.  
25

44. A composition comprising a nucleic acid molecule of any of Claims 1, 2, or 3 and a pharmaceutically acceptable formulation agent.

30 45. The composition of Claim 44, wherein said nucleic acid molecule is contained in a viral vector.

46. A viral vector comprising a nucleic acid molecule of any of Claims 1, 2, or 3.

47. A fusion polypeptide comprising the polypeptide of any of Claims 13, 14, or 15 fused to a heterologous amino acid sequence.

48. The fusion polypeptide of Claim 47, wherein the heterologous amino acid sequence is an IgG constant domain or fragment thereof.

49. A method for treating, preventing, or ameliorating a medical condition comprising administering to a patient the polypeptide of any of Claims 13, 14, or 15, or the polypeptide encoded by the nucleic acid of any of Claims 1, 2, or 3.

50. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or amount of expression of the polypeptide of any of Claims 13, 14, or 15, or the polypeptide encoded by the nucleic acid molecule of any of Claims 1, 2, or 3 in a sample; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

51. A device, comprising:

(a) a membrane suitable for implantation; and

(b) cells encapsulated within said membrane, wherein said cells secrete a protein of any of Claims 13, 14, or 15; and

said membrane is permeable to said protein and impermeable to materials detrimental to said cells.

52. A method of identifying a compound that binds to an ABCL polypeptide comprising:

- (a) contacting the polypeptide of any of Claims 13, 14, or 15 with a compound; and
- (b) determining the extent of binding of the ABCL polypeptide to the compound.

5

53. The method of Claim 52, further comprising determining the activity of the polypeptide when bound to the compound.

54. A method of modulating levels of a polypeptide in an animal comprising administering to the animal the nucleic acid molecule of any of Claims 1, 2, or 3.

55. A transgenic non-human mammal comprising the nucleic acid molecule of any of Claims 1, 2, or 3.

15

56. A process for determining whether a compound inhibits ABCL polypeptide activity or ABCL polypeptide production comprising exposing a transgenic mammal according to Claim 55 to the compound, and measuring ABCL polypeptide activity or ABCL polypeptide production in said mammal.

20

57. A nucleic acid molecule of any of Claims 1, 2, or 3 attached to a solid support.

58. An array of nucleic acid molecules comprising at least one nucleic acid molecule of any of Claims 1, 2, or 3.

25

## FIG. 1A

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| atg gct ttc tgc aca cag ttg atg ctt ctg ctg tgg aaa aat tac acc | 48  |
| Met Ala Phe Cys Thr Gln Leu Met Leu Leu Leu Trp Lys Asn Tyr Thr |     |
| 1 5 10 15                                                       |     |
| tat cga cgg aga caa ccg atc caa cta cta gtg gag ttg ctt tgg ccc | 96  |
| Tyr Arg Arg Arg Gln Pro Ile Gln Leu Leu Val Glu Leu Leu Trp Pro |     |
| 20 25 30                                                        |     |
| ctc ttc ctc ttc ttc atc cta gtg gct gtc cgt cac tcc cac ccc cct | 144 |
| Leu Phe Leu Phe Phe Ile Leu Val Ala Val Arg His Ser His Pro Pro |     |
| 35 40 45                                                        |     |
| ctg gag cat cac gaa tgc cac ttt cca aac aag cca tta cca tcg gcg | 192 |
| Leu Glu His His Glu Cys His Phe Pro Asn Lys Pro Leu Pro Ser Ala |     |
| 50 55 60                                                        |     |
| ggc acg gtg ccc tgg ctg cag ggc ctt gtc tgc aac gta aac aac tcc | 240 |
| Gly Thr Val Pro Trp Leu Gln Gly Leu Val Cys Asn Val Asn Asn Ser |     |
| 65 70 75 80                                                     |     |
| tgc ttc cag cac cca acg cct ggc gag aag cct ggg gtc ctg agt aac | 288 |
| Cys Phe Gln His Pro Thr Pro Gly Glu Lys Pro Gly Val Leu Ser Asn |     |
| 85 90 95                                                        |     |
| ttt aag gat tcc ttg atc tcg agg ctc ctc gct gat acc cgc aca gtg | 336 |
| Phe Lys Asp Ser Leu Ile Ser Arg Leu Leu Ala Asp Thr Arg Thr Val |     |
| 100 105 110                                                     |     |
| ctc ggg ggc cac agc atc cag gac atg ctg gat gcc ctg ggg aaa ctg | 384 |
| Leu Gly Gly His Ser Ile Gln Asp Met Leu Asp Ala Leu Gly Lys Leu |     |
| 115 120 125                                                     |     |
| atc ccc gtg ctc agg gca gtt gga ggt gga gca cga cca cag gag agt | 432 |
| Ile Pro Val Leu Arg Ala Val Gly Gly Gly Ala Arg Pro Gln Glu Ser |     |
| 130 135 140                                                     |     |
| gac cag ccg acc agt caa ggg tca gtg act aag ctt ctg gag aag atc | 480 |
| Asp Gln Pro Thr Ser Gln Gly Ser Val Thr Lys Leu Leu Glu Lys Ile |     |
| 145 150 155 160                                                 |     |
| ctg caa agg gca tcc ctg gat cct gtg ctg ggt caa gcc cag gat tct | 528 |
| Leu Gln Arg Ala Ser Leu Asp Pro Val Leu Gly Gln Ala Gln Asp Ser |     |
| 165 170 175                                                     |     |
| atg aga aag ttc tca gat gct atc agg gat ctt gcc cag gag ctc ctg | 576 |
| Met Arg Lys Phe Ser Asp Ala Ile Arg Asp Leu Ala Gln Glu Leu Leu |     |
| 180 185 190                                                     |     |
| aca ctg ccc agc ctg atg gag ctc cga gct ttg ctg cgg agg ccc cga | 624 |
| Thr Leu Pro Ser Leu Met Glu Leu Arg Ala Leu Leu Arg Pro Arg     |     |
| 195 200 205                                                     |     |

## FIG. 1B

|                                                                 |      |
|-----------------------------------------------------------------|------|
| ggg tca gct ggt tct ctg gag ctg gtt tcg gag gcc ctc tgc agt acc | 672  |
| Gly Ser Ala Gly Ser Leu Glu Leu Val Ser Glu Ala Leu Cys Ser Thr |      |
| 210 215 220                                                     |      |
| aag gga ccc agc agt cca ggg ggc ctg tcc ctc aat tgg tac gaa gcc | 720  |
| Lys Gly Pro Ser Ser Pro Gly Gly Leu Ser Leu Asn Trp Tyr Glu Ala |      |
| 225 230 235 240                                                 |      |
| aac cag ctt aat gag ttc atg ggg cca gag gtg gcc cct gcc ctg cct | 768  |
| Asn Gln Leu Asn Glu Phe Met Gly Pro Glu Val Ala Pro Ala Leu Pro |      |
| 245 250 255                                                     |      |
| gac aac agt ctc agc cct gcc tgc tct gag ttt gtg ggg aca ctg gat | 816  |
| Asp Asn Ser Leu Ser Pro Ala Cys Ser Glu Phe Val Gly Thr Leu Asp |      |
| 260 265 270                                                     |      |
| gac cac cct gtg tct cgg ctg ctc tgg agg cgc ctg aag cca ttg atc | 864  |
| Asp His Pro Val Ser Arg Leu Leu Trp Arg Arg Leu Lys Pro Leu Ile |      |
| 275 280 285                                                     |      |
| ctc ggg aaa att ctc ttt gca cct gac aca aac ttc act cgg aag ctc | 912  |
| Leu Gly Lys Ile Leu Phe Ala Pro Asp Thr Asn Phe Thr Arg Lys Leu |      |
| 290 295 300                                                     |      |
| atg gct cag gtg aac cag acc ttc gag gag ctg gct ctg ttg agg gac | 960  |
| Met Ala Gln Val Asn Gln Thr Phe Glu Glu Leu Ala Leu Leu Arg Asp |      |
| 305 310 315 320                                                 |      |
| cta cac gaa ctc tgg ggg gtg ctg gga ccc cag atc ttc aac ttc atg | 1008 |
| Leu His Glu Leu Trp Gly Val Leu Gly Pro Gln Ile Phe Asn Phe Met |      |
| 325 330 335                                                     |      |
| aat gac agt acc aac gtg gcc atg ctt cag agg ctt ctg gat gtg ggg | 1056 |
| Asn Asp Ser Thr Asn Val Ala Met Leu Gln Arg Leu Leu Asp Val Gly |      |
| 340 345 350                                                     |      |
| ggc aca ggg cag agg cag cag aca ccc aga gcc cag aag aag ttg gag | 1104 |
| Gly Thr Gly Gln Arg Gln Gln Thr Pro Arg Ala Gln Lys Lys Leu Glu |      |
| 355 360 365                                                     |      |
| gct atc aaa gac ttt ctg gat cct agt agg ggt ggc tac agc tgg cgg | 1152 |
| Ala Ile Lys Asp Phe Leu Asp Pro Ser Arg Gly Gly Tyr Ser Trp Arg |      |
| 370 375 380                                                     |      |
| gag gcc cac gca gat atg gga cgc ctg gct gga atc cta gga caa atg | 1200 |
| Glu Ala His Ala Asp Met Gly Arg Leu Ala Gly Ile Leu Gly Gln Met |      |
| 385 390 395 400                                                 |      |
| atg gag tgt gtg tcc ctg gac aag ctg gag gct gtg ccc tca gag gaa | 1248 |
| Met Glu Cys Val Ser Leu Asp Lys Leu Glu Ala Val Pro Ser Glu Glu |      |
| 405 410 415                                                     |      |

## FIG. 1C

|                                                                 |      |
|-----------------------------------------------------------------|------|
| gct ctt gtg tcc cgt gcc ctg gag ctg ctg ggt gag cgc cgc ctc tgg | 1296 |
| Ala Leu Val Ser Arg Ala Leu Glu Leu Leu Gly Glu Arg Arg Leu Trp |      |
| 420 425 430                                                     |      |
| gca ggc atc gtg ttc ctg agc cca gag cat cct ctg gac cca tcc gaa | 1344 |
| Ala Gly Ile Val Phe Leu Ser Pro Glu His Pro Leu Asp Pro Ser Glu |      |
| 435 440 445                                                     |      |
| ctg tca tct cca gcc ctg agt cct ggc cac cta cga ttc aag att cga | 1392 |
| Leu Ser Ser Pro Ala Leu Ser Pro Gly His Leu Arg Phe Lys Ile Arg |      |
| 450 455 460                                                     |      |
| atg gat atc gat gat gtc aca agg acc aat aag atc agg gac aag ttt | 1440 |
| Met Asp Ile Asp Asp Val Thr Arg Thr Asn Lys Ile Arg Asp Lys Phe |      |
| 465 470 475 480                                                 |      |
| tgg gac cca ggt ccg tca gca gat cct ttc atg gac ctt cgg tat gtg | 1488 |
| Trp Asp Pro Gly Pro Ser Ala Asp Pro Phe Met Asp Leu Arg Tyr Val |      |
| 485 490 495                                                     |      |
| tgg gga ggc ttc gtg tac ctg cag gac ctg ctg gag cag gca gct gtg | 1536 |
| Trp Gly Gly Phe Val Tyr Leu Gln Asp Leu Leu Glu Gln Ala Ala Val |      |
| 500 505 510                                                     |      |
| cga gtg ctc ggt ggc ggg aac tcc cgc aca ggt ctc tac ctg cag cag | 1584 |
| Arg Val Leu Gly Gly Gly Asn Ser Arg Thr Gly Leu Tyr Leu Gln Gln |      |
| 515 520 525                                                     |      |
| atg cca cac ccc tgc tac gtg gat gat gtg ttc ctg cgg gtg ctg agc | 1632 |
| Met Pro His Pro Cys Tyr Val Asp Asp Val Phe Leu Arg Val Leu Ser |      |
| 530 535 540                                                     |      |
| cgg tct ctg cct ctg ttt ctg act ctg gcc tgg att tat tcg gtg gcg | 1680 |
| Arg Ser Leu Pro Leu Phe Leu Thr Leu Ala Trp Ile Tyr Ser Val Ala |      |
| 545 550 555 560                                                 |      |
| ctc act gtg aag gcc gtg gtg cgt gag aaa gag aca cgg ctg cga gaa | 1728 |
| Leu Thr Val Lys Ala Val Val Arg Glu Lys Glu Thr Arg Leu Arg Glu |      |
| 565 570 575                                                     |      |
| acc atg cgt gcg atg ggg ctg agc cgc gcg gtg ctc tgg ctt ggt tgg | 1776 |
| Thr Met Arg Ala Met Gly Leu Ser Arg Ala Val Leu Trp Leu Gly Trp |      |
| 580 585 590                                                     |      |
| ttc ctc agc tgc ctg gga ccc ttc ctg gtc agc gct gcg ttg ctg gta | 1824 |
| Phe Leu Ser Cys Leu Gly Pro Phe Leu Val Ser Ala Ala Leu Leu Val |      |
| 595 600 605                                                     |      |
| tta gtg ctt aag cta ggg aac atc ctt cct tac agc cac ccg gtt gta | 1872 |
| Leu Val Leu Lys Leu Gly Asn Ile Leu Pro Tyr Ser His Pro Val Val |      |
| 610 615 620                                                     |      |

## FIG. 1D

|                                                                 |      |
|-----------------------------------------------------------------|------|
| atc ttc ctt ttc ttg gcg gcc ttc gcg gtg gcc acc gtc gct cag agt | 1920 |
| Ile Phe Leu Phe Leu Ala Ala Phe Ala Val Ala Thr Val Ala Gln Ser |      |
| 625 630 635 640                                                 |      |
| ttt ctg ctc agc gcc ttc ttc tcc agg gcc aat ctg gca gca gcc tgc | 1968 |
| Phe Leu Leu Ser Ala Phe Phe Ser Arg Ala Asn Leu Ala Ala Ala Cys |      |
| 645 650 655                                                     |      |
| ggg ggg ctc gcc tat ttt gcg ctc tat ctg ccc tac gta ctg tgt gtg | 2016 |
| Gly Gly Leu Ala Tyr Phe Ala Leu Tyr Leu Pro Tyr Val Leu Cys Val |      |
| 660 665 670                                                     |      |
| gcc tgg cgc gag cgc ctg cac ctg ggc gga ctc tta gct gcg agc ctg | 2064 |
| Ala Trp Arg Glu Arg Leu His Leu Gly Gly Leu Leu Ala Ala Ser Leu |      |
| 675 680 685                                                     |      |
| ctg tcc cct gta gcc ttt ggc ttt gga tgc gaa agc ctg gcg cta cta | 2112 |
| Leu Ser Pro Val Ala Phe Gly Phe Gly Cys Glu Ser Leu Ala Leu Leu |      |
| 690 695 700                                                     |      |
| gag gag cag gga gac ggg gct cag tgg cac aat ttg ggc aca ggc ccc | 2160 |
| Glu Glu Gln Gly Asp Gly Ala Gln Trp His Asn Leu Gly Thr Gly Pro |      |
| 705 710 715 720                                                 |      |
| gcg gag gac gtc ttc agc ctg gcc cag gtg tct gcc ttc ctg ttg ctt | 2208 |
| Ala Glu Asp Val Phe Ser Leu Ala Gln Val Ser Ala Phe Leu Leu Leu |      |
| 725 730 735                                                     |      |
| gat gcc gtc atc tac ggc ctt gcc ctc tgg tac cta gag gct gtg tgc | 2256 |
| Asp Ala Val Ile Tyr Gly Leu Ala Leu Trp Tyr Leu Glu Ala Val Cys |      |
| 740 745 750                                                     |      |
| cca ggc cag tat gga atc cct gaa cca tgg aat ttc cct ttt cgg agg | 2304 |
| Pro Gly Gln Tyr Gly Ile Pro Glu Pro Trp Asn Phe Pro Phe Arg Arg |      |
| 755 760 765                                                     |      |
| agc tac tgg tgt gga cct ggg cct ccc aag agt tct gtc ttg gcc cct | 2352 |
| Ser Tyr Trp Cys Gly Pro Gly Pro Pro Lys Ser Ser Val Leu Ala Pro |      |
| 770 775 780                                                     |      |
| gcc cca caa gat ccc aag gtt ctg gtg gaa gag cca ccc ctt ggc ctg | 2400 |
| Ala Pro Gln Asp Pro Lys Val Leu Val Glu Glu Pro Pro Leu Gly Leu |      |
| 785 790 795 800                                                 |      |
| gtt cct ggt gtc tcc att cga ggc ctg aag aaa cat ttt cgt ggc tgt | 2448 |
| Val Pro Gly Val Ser Ile Arg Gly Leu Lys Lys His Phe Arg Gly Cys |      |
| 805 810 815                                                     |      |
| ccg cag cca gcc ctg caa gga ctc aac ctt gac ttc tac gaa ggc cac | 2496 |
| Pro Gln Pro Ala Leu Gln Gly Leu Asn Leu Asp Phe Tyr Glu Gly His |      |
| 820 825 830                                                     |      |

## FIG. 1E

|                                                                 |      |
|-----------------------------------------------------------------|------|
| atc act gcc ttt ttg ggt cac aac ggg gct ggc aag aca acc aca ctg | 2544 |
| Ile Thr Ala Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu |      |
| 835 840 845                                                     |      |
| tcc att ttg agt ggt ctc ttc cca ccc agt agt ggc tcg gcc tcc atc | 2592 |
| Ser Ile Leu Ser Gly Leu Phe Pro Pro Ser Ser Gly Ser Ala Ser Ile |      |
| 850 855 860                                                     |      |
| ctg ggc cat gat gta caa acc aac atg gca gcc atc cgg ccc cac ctg | 2640 |
| Leu Gly His Asp Val Gln Thr Asn Met Ala Ala Ile Arg Pro His Leu |      |
| 865 870 875 880                                                 |      |
| ggc atc tgc ccg cag tac aat gtg ctg ttt gat atg ctg aca gtg gaa | 2688 |
| Gly Ile Cys Pro Gln Tyr Asn Val Leu Phe Asp Met Leu Thr Val Glu |      |
| 885 890 895                                                     |      |
| gaa cat gtt tgg ttc tat ggc cgt ttg aaa ggc gtg agt gca gcc gcc | 2736 |
| Glu His Val Trp Phe Tyr Gly Arg Leu Lys Gly Val Ser Ala Ala Ala |      |
| 900 905 910                                                     |      |
| atg ggc ccc gag cgg gaa cgt ctg ata cgg gat gtg ggg ctt acc ctc | 2784 |
| Met Gly Pro Glu Arg Glu Arg Leu Ile Arg Asp Val Gly Leu Thr Leu |      |
| 915 920 925                                                     |      |
| aag cgg gac aca cag aca cgc cac ctc tct ggt gga atg cag aga aaa | 2832 |
| Lys Arg Asp Thr Gln Thr Arg His Leu Ser Gly Gly Met Gln Arg Lys |      |
| 930 935 940                                                     |      |
| ctt tct gtg gcc att gcc ttt gtg ggt ggc tct cgt gtg gtc atc atg | 2880 |
| Leu Ser Val Ala Ile Ala Phe Val Gly Gly Ser Arg Val Val Ile Met |      |
| 945 950 955 960                                                 |      |
| gac gag ccc act gct ggt gtg gac ccc gct tcc cgc cgt ggc att tgg | 2928 |
| Asp Glu Pro Thr Ala Gly Val Asp Pro Ala Ser Arg Arg Gly Ile Trp |      |
| 965 970 975                                                     |      |
| gaa ttg cta ctt aag tac aga gaa ggt cgg aca ctg att ctc tcc act | 2976 |
| Glu Leu Leu Leu Lys Tyr Arg Glu Gly Arg Thr Leu Ile Leu Ser Thr |      |
| 980 985 990                                                     |      |
| cac cac ctg gat gag gca gag ctc ttg gga gat cgc gtg gcc atg gtg | 3024 |
| His His Leu Asp Glu Ala Glu Leu Leu Gly Asp Arg Val Ala Met Val |      |
| 995 1000 1005                                                   |      |
| gca ggt ggc tct ttg tgc tgc tgt ggg tcc ccg ctt ttc ttg cgc cga | 3072 |
| Ala Gly Gly Ser Leu Cys Cys Cys Gly Ser Pro Leu Phe Leu Arg Arg |      |
| 1010 1015 1020                                                  |      |
| cac ttg ggc tgc ggt tac tac ctg acc ctg gtg aag agt tct cag tcc | 3120 |
| His Leu Gly Cys Gly Tyr Tyr Leu Thr Leu Val Lys Ser Ser Gln Ser |      |
| 1025 1030 1035 1040                                             |      |

## FIG. 1F

|                                                                 |      |
|-----------------------------------------------------------------|------|
| ctc gtc acc cat gat gct aag gga gac agt gag gac ccc aga cgg gaa | 3168 |
| Leu Val Thr His Asp Ala Lys Gly Asp Ser Glu Asp Pro Arg Arg Glu |      |
| 1045 1050 1055                                                  |      |
| aag aag tca gat ggc aat ggc agg acg tca gac aca gcg ttc aca cga | 3216 |
| Lys Lys Ser Asp Gly Asn Gly Arg Thr Ser Asp Thr Ala Phe Thr Arg |      |
| 1060 1065 1070                                                  |      |
| gga acc tca gac aag agc aac cag gcc ccg gct cct ggc gcc gtt ccc | 3264 |
| Gly Thr Ser Asp Lys Ser Asn Gln Ala Pro Ala Pro Gly Ala Val Pro |      |
| 1075 1080 1085                                                  |      |
| atc acc cca agc aca gcc cgg ata cta gag cta gtg cag cag cat gtg | 3312 |
| Ile Thr Pro Ser Thr Ala Arg Ile Leu Glu Leu Val Gln Gln His Val |      |
| 1090 1095 1100                                                  |      |
| cct gga gca caa ctc gtg gag gac ctg ccc cat gag ctt ctg ctt gtg | 3360 |
| Pro Gly Ala Gln Leu Val Glu Asp Leu Pro His Glu Leu Leu Leu Val |      |
| 1105 1110 1115 1120                                             |      |
| cta ccc tat gcg ggg gcc ctg gat ggc agc ttc gcc atg gtc ttc cag | 3408 |
| Leu Pro Tyr Ala Gly Ala Leu Asp Gly Ser Phe Ala Met Val Phe Gln |      |
| 1125 1130 1135                                                  |      |
| gag ctg gat cag cag ctg gag ctc ctg ggt ctc aca ggc tac ggg atc | 3456 |
| Glu Leu Asp Gln Gln Leu Glu Leu Leu Gly Leu Thr Gly Tyr Gly Ile |      |
| 1140 1145 1150                                                  |      |
| tcg gac acc aac ctg gag gag atc ttc cta aag gtg gtg gag gat gcg | 3504 |
| Ser Asp Thr Asn Leu Glu Glu Ile Phe Leu Lys Val Val Glu Asp Ala |      |
| 1155 1160 1165                                                  |      |
| cac aga gaa ggt ggg gac tct aga ccg cag ctg cac ctt cgc aca tgc | 3552 |
| His Arg Glu Gly Gly Asp Ser Arg Pro Gln Leu His Leu Arg Thr Cys |      |
| 1170 1175 1180                                                  |      |
| act cca cag ccc ccc aca ggg cca gag gca tca gtt ctg gag aat ggg | 3600 |
| Thr Pro Gln Pro Pro Thr Gly Pro Glu Ala Ser Val Leu Glu Asn Gly |      |
| 1185 1190 1195 1200                                             |      |
| gag ctg gct aag ctg gtg ctg gat ccc caa gcc cca cag ggc ttg gca | 3648 |
| Glu Leu Ala Lys Leu Val Leu Asp Pro Gln Ala Pro Gln Gly Leu Ala |      |
| 1205 1210 1215                                                  |      |
| ccc aac gct gcc caa gtg caa ggc tgg aca ctt acc tgt caa cag ctc | 3696 |
| Pro Asn Ala Ala Gln Val Gln Gly Trp Thr Leu Thr Cys Gln Gln Leu |      |
| 1220 1225 1230                                                  |      |
| cgg gct ctg ctc cac aag cgt ttt ctg ctt gct cgc cgc agc cgc cgg | 3744 |
| Arg Ala Leu Leu His Lys Arg Phe Leu Leu Ala Arg Arg Ser Arg Arg |      |
| 1235 1240 1245                                                  |      |

## FIG. 1G

|                                                                 |      |
|-----------------------------------------------------------------|------|
| ggc ctg ttt gca cag gtt gtg ttg cct gcc ctc ttt gtg ggc ctg gcc | 3792 |
| Gly Leu Phe Ala Gln Val Val Leu Pro Ala Leu Phe Val Gly Leu Ala |      |
| 1250 1255 1260                                                  |      |
| ctg ttc ttc agc ctc att gtg cct cct ttt ggc cag tac cca ccc ctg | 3840 |
| Leu Phe Phe Ser Leu Ile Val Pro Pro Phe Gly Gln Tyr Pro Pro Leu |      |
| 1265 1270 1275 1280                                             |      |
| cag ctc agc cct gct atg tat ggc cct cag gtc tcg ttc ttc agt gag | 3888 |
| Gln Leu Ser Pro Ala Met Tyr Gly Pro Gln Val Ser Phe Phe Ser Glu |      |
| 1285 1290 1295                                                  |      |
| gat gcc cct ggg gac ccc aac cgg atg aag ctg ctg gag gct ctg cta | 3936 |
| Asp Ala Pro Gly Asp Pro Asn Arg Met Lys Leu Leu Glu Ala Leu Leu |      |
| 1300 1305 1310                                                  |      |
| ggg gag gct ggg ctg cag gaa ccc agt atg cag gac aaa gat gcc agg | 3984 |
| Gly Glu Ala Gly Leu Gln Glu Pro Ser Met Gln Asp Lys Asp Ala Arg |      |
| 1315 1320 1325                                                  |      |
| gga tct gag tgt aca cac tcc cta gct tgc tac ttc acg gtc cct gag | 4032 |
| Gly Ser Glu Cys Thr His Ser Leu Ala Cys Tyr Phe Thr Val Pro Glu |      |
| 1330 1335 1340                                                  |      |
| gtc cct cct gat gtg gcc agc atc ctg gcc agt ggc aac tgg acg cca | 4080 |
| Val Pro Pro Asp Val Ala Ser Ile Leu Ala Ser Gly Asn Trp Thr Pro |      |
| 1345 1350 1355 1360                                             |      |
| gaa tct cca tcc cca gct tgc caa tgc agt cag cct gga gcc cgc cgc | 4128 |
| Glu Ser Pro Ser Pro Ala Cys Gln Cys Ser Gln Pro Gly Ala Arg Arg |      |
| 1365 1370 1375                                                  |      |
| ctg ttg cca gat tgc ccg gct gga gct ggg ggt cca cca ccc ccc cag | 4176 |
| Leu Leu Pro Asp Cys Pro Ala Gly Ala Gly Gly Pro Pro Pro Gln     |      |
| 1380 1385 1390                                                  |      |
| gct gtg gct ggc ttg ggg gag gtg gtc cag aac ctc act ggc cga aat | 4224 |
| Ala Val Ala Gly Leu Gly Glu Val Val Gln Asn Leu Thr Gly Arg Asn |      |
| 1395 1400 1405                                                  |      |
| gtg tct gac ttt ttg gtg aag aca tac ccc agc ctg gtg cgc cga ggc | 4272 |
| Val Ser Asp Phe Leu Val Lys Thr Tyr Pro Ser Leu Val Arg Arg Gly |      |
| 1410 1415 1420                                                  |      |
| cta aag acc aag aag tgg gtg gat gag gtc aga tat ggg ggc ttc tcc | 4320 |
| Leu Lys Thr Lys Lys Trp Val Asp Glu Val Arg Tyr Gly Gly Phe Ser |      |
| 1425 1430 1435 1440                                             |      |
| ctg gga ggc cga gat cca gac ctg ccc aca ggg cat gag gtg gtc cgc | 4368 |
| Leu Gly Gly Arg Asp Pro Asp Leu Pro Thr Gly His Glu Val Val Arg |      |
| 1445 1450 1455                                                  |      |

## FIG. 1H

|                                                                 |      |
|-----------------------------------------------------------------|------|
| aca ttg gca gag att cgg gca ctg ctg agc ccc caa cct ggg aat gcg | 4416 |
| Thr Leu Ala Glu Ile Arg Ala Leu Leu Ser Pro Gln Pro Gly Asn Ala |      |
| 1460 1465 1470                                                  |      |
| cta gac cgt atc ctg aac aac ctc act cag tgg gcc ctt ggc ctt gat | 4464 |
| Leu Asp Arg Ile Leu Asn Asn Leu Thr Gln Trp Ala Leu Gly Leu Asp |      |
| 1475 1480 1485                                                  |      |
| gct cgg aac agc ctc aag atc tgg ttc aac aac aag ggc tgg cat gcc | 4512 |
| Ala Arg Asn Ser Leu Lys Ile Trp Phe Asn Asn Lys Gly Trp His Ala |      |
| 1490 1495 1500                                                  |      |
| atg gtg gcc ttt gtg aac cga gcc aac aat gga ctc cta cat gcc ctc | 4560 |
| Met Val Ala Phe Val Asn Arg Ala Asn Asn Gly Leu Leu His Ala Leu |      |
| 1505 1510 1515 1520                                             |      |
| cta cca tct ggt cca gtc cgc cat gcc cac agc atc act aca ctc aac | 4608 |
| Leu Pro Ser Gly Pro Val Arg His Ala His Ser Ile Thr Thr Leu Asn |      |
| 1525 1530 1535                                                  |      |
| cat cct ttg aac ttg acc aag gag cag cta tct gaa gct aca ctg ata | 4656 |
| His Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ala Thr Leu Ile |      |
| 1540 1545 1550                                                  |      |
| gcc tcc tct gtg gat gtc ctt gtc tcc atc tgt gtg gtc ttc gcc atg | 4704 |
| Ala Ser Ser Val Asp Val Leu Val Ser Ile Cys Val Val Phe Ala Met |      |
| 1555 1560 1565                                                  |      |
| tca ttt gtc cca gcc agc ttt acc ctg gtc ctc ata gag gaa cgc atc | 4752 |
| Ser Phe Val Pro Ala Ser Phe Thr Leu Val Leu Ile Glu Glu Arg Ile |      |
| 1570 1575 1580                                                  |      |
| acc aga gcc aag cat ctg cag ctg gtc agc ggc ctg ccc caa acc ctc | 4800 |
| Thr Arg Ala Lys His Leu Gln Leu Val Ser Gly Leu Pro Gln Thr Leu |      |
| 1585 1590 1595 1600                                             |      |
| tat tgg ctt ggc aac ttc ctc tgg gac atg tgt aac tac ttg gtg gca | 4848 |
| Tyr Trp Leu Gly Asn Phe Leu Trp Asp Met Cys Asn Tyr Leu Val Ala |      |
| 1605 1610 1615                                                  |      |
| gtg tgc ata gtg gtg ttc atc ttc cta gcc ttt cag cag aga gcc tat | 4896 |
| Val Cys Ile Val Val Phe Ile Phe Leu Ala Phe Gln Gln Arg Ala Tyr |      |
| 1620 1625 1630                                                  |      |
| gtg gcc cca gag aac ctg cct gct ctc tta ctc ttg ctt ctg ctg tat | 4944 |
| Val Ala Pro Glu Asn Leu Pro Ala Leu Leu Leu Leu Leu Leu Tyr     |      |
| 1635 1640 1645                                                  |      |
| ggg tgg tct atc aca cca ctc atg tac cca gcc tcc ttc ttc ttc tca | 4992 |
| Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala Ser Phe Phe Phe Ser |      |
| 1650 1655 1660                                                  |      |

## FIG. 1I

|      |     |     |      |     |     |      |     |     |      |     |     |     |     |     |     |      |
|------|-----|-----|------|-----|-----|------|-----|-----|------|-----|-----|-----|-----|-----|-----|------|
| gtg  | ccc | agc | acg  | gcc | tat | gtg  | gtg | ctc | acc  | tgc | atc | aac | ctc | ttc | att | 5040 |
| Val  | Pro | Ser | Thr  | Ala | Tyr | Val  | Val | Leu | Thr  | Cys | Ile | Asn | Leu | Phe | Ile |      |
| 1665 |     |     | 1670 |     |     | 1675 |     |     | 1680 |     |     |     |     |     |     |      |
| ggc  | atc | aat | agc  | agc | atg | gcc  | acc | ttc | gtg  | cta | gaa | ctg | ctt | tca | gat | 5088 |
| Gly  | Ile | Asn | Ser  | Ser | Met | Ala  | Thr | Phe | Val  | Leu | Glu | Leu | Leu | Ser | Asp |      |
| 1685 |     |     | 1690 |     |     | 1695 |     |     |      |     |     |     |     |     |     |      |
| cag  | aac | ctg | caa  | gaa | gtg | agc  | cgg | atc | ctg  | aaa | caa | gtg | ttt | ctt | att | 5136 |
| Gln  | Asn | Leu | Gln  | Glu | Val | Ser  | Arg | Ile | Leu  | Lys | Gln | Val | Phe | Leu | Ile |      |
| 1700 |     |     | 1705 |     |     | 1710 |     |     |      |     |     |     |     |     |     |      |
| ttc  | ccc | cac | ttt  | tgc | ctt | ggc  | cga | ggg | ctc  | att | gac | atg | gtt | cgg | aac | 5184 |
| Phe  | Pro | His | Phe  | Cys | Leu | Gly  | Arg | Gly | Leu  | Ile | Asp | Met | Val | Arg | Asn |      |
| 1715 |     |     | 1720 |     |     | 1725 |     |     |      |     |     |     |     |     |     |      |
| cag  | gcc | atg | gca  | gat | gcc | ttt  | gag | cgc | tta  | gga | gac | aag | caa | ttt | cag | 5232 |
| Gln  | Ala | Met | Ala  | Asp | Ala | Phe  | Glu | Arg | Leu  | Gly | Asp | Lys | Gln | Phe | Gln |      |
| 1730 |     |     | 1735 |     |     | 1740 |     |     |      |     |     |     |     |     |     |      |
| tca  | ccc | cta | cgc  | tgg | gac | atc  | att | ggc | aag  | aac | ctc | ctg | gcc | atg | atg | 5280 |
| Ser  | Pro | Leu | Arg  | Trp | Asp | Ile  | Ile | Gly | Lys  | Asn | Leu | Leu | Ala | Met | Met |      |
| 1745 |     |     | 1750 |     |     | 1755 |     |     | 1760 |     |     |     |     |     |     |      |
| gcc  | cag | gga | cct  | ctg | ttc | ctc  | ctc | atc | aca  | ctc | ctg | ctc | caa | cac | cgc | 5328 |
| Ala  | Gln | Gly | Pro  | Leu | Phe | Leu  | Leu | Ile | Thr  | Leu | Leu | Leu | Gln | His | Arg |      |
| 1765 |     |     | 1770 |     |     | 1775 |     |     |      |     |     |     |     |     |     |      |
| aac  | cgt | ctc | ctg  | cca | caa | tca  | aaa | cca | aga  | ctg | ctg | ccg | ccc | ctg | ggg | 5376 |
| Asn  | Arg | Leu | Leu  | Pro | Gln | Ser  | Lys | Pro | Arg  | Leu | Leu | Pro | Pro | Leu | Gly |      |
| 1780 |     |     | 1785 |     |     | 1790 |     |     |      |     |     |     |     |     |     |      |
| gag  | gag | gat | gag  | gat | gtg | gct  | caa | gag | cgt  | gag | cgg | gtg | acc | aag | ggg | 5424 |
| Glu  | Glu | Asp | Glu  | Asp | Val | Ala  | Gln | Glu | Arg  | Glu | Arg | Val | Thr | Lys | Gly |      |
| 1795 |     |     | 1800 |     |     | 1805 |     |     |      |     |     |     |     |     |     |      |
| gct  | acc | cag | ggg  | gat | gtg | cta  | gtc | ctc | agg  | gac | ttg | acc | aag | gtt | tac | 5472 |
| Ala  | Thr | Gln | Gly  | Asp | Val | Leu  | Val | Leu | Arg  | Asp | Leu | Thr | Lys | Val | Tyr |      |
| 1810 |     |     | 1815 |     |     | 1820 |     |     |      |     |     |     |     |     |     |      |
| cgt  | ggg | cag | agg  | aac | cca | gct  | gtg | gat | cgc  | ctg | tgc | tta | ggg | atc | ccc | 5520 |
| Arg  | Gly | Gln | Arg  | Asn | Pro | Ala  | Val | Asp | Arg  | Leu | Cys | Leu | Gly | Ile | Pro |      |
| 1825 |     |     | 1830 |     |     | 1835 |     |     | 1840 |     |     |     |     |     |     |      |
| cct  | ggg | gag | tgt  | ttc | ggg | ctg  | ctg | ggt | gtc  | aac | ggg | gca | ggg | aag | aca | 5568 |
| Pro  | Gly | Glu | Cys  | Phe | Gly | Leu  | Leu | Gly | Val  | Asn | Gly | Ala | Gly | Lys | Thr |      |
| 1845 |     |     | 1850 |     |     | 1855 |     |     |      |     |     |     |     |     |     |      |
| tcc  | acc | ttc | cgc  | atg | gtg | aca  | ggg | gac | aca  | ctg | ccc | agc | agt | ggg | gaa | 5616 |
| Ser  | Thr | Phe | Arg  | Met | Val | Thr  | Gly | Asp | Thr  | Leu | Pro | Ser | Ser | Gly | Glu |      |
| 1860 |     |     | 1865 |     |     | 1870 |     |     |      |     |     |     |     |     |     |      |

## FIG. 1J

|                                                                 |      |
|-----------------------------------------------------------------|------|
| gca gta ctg gca ggc cac aac gtg gcc cag gag cgg tct gcc gca cac | 5664 |
| Ala Val Leu Ala Gly His Asn Val Ala Gln Glu Arg Ser Ala Ala His |      |
| 1875 1880 1885                                                  |      |
| cgc agc atg ggc tac tgt ccc cag tct gat gcc atc ttc gac ctg ctg | 5712 |
| Arg Ser Met Gly Tyr Cys Pro Gln Ser Asp Ala Ile Phe Asp Leu Leu |      |
| 1890 1895 1900                                                  |      |
| acc ggc cgg gaa cat ctg gaa ctg ttt gct cgc ctg cgc ggg gtg ccc | 5760 |
| Thr Gly Arg Glu His Leu Glu Leu Phe Ala Arg Leu Arg Gly Val Pro |      |
| 1905 1910 1915 1920                                             |      |
| gag gcc caa gtt gcc cag act gcg ctc tct ggc ctg gtg cgc ctg ggc | 5808 |
| Glu Ala Gln Val Ala Gln Thr Ala Leu Ser Gly Leu Val Arg Leu Gly |      |
| 1925 1930 1935                                                  |      |
| ctt cct agc tat gca gac cga ccc gcg ggt acc tac agc gga ggc aac | 5856 |
| Leu Pro Ser Tyr Ala Asp Arg Pro Ala Gly Thr Tyr Ser Gly Gly Asn |      |
| 1940 1945 1950                                                  |      |
| aaa cgg aag ctg gcg aca gcc tta gct ctg gtt ggt gac cca gct gtg | 5904 |
| Lys Arg Lys Leu Ala Thr Ala Leu Ala Leu Val Gly Asp Pro Ala Val |      |
| 1955 1960 1965                                                  |      |
| gtc ttt ctg gac gag ccc acc aca ggc atg gac cca agt gcg cgg cga | 5952 |
| Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Ser Ala Arg Arg |      |
| 1970 1975 1980                                                  |      |
| ttt ctt tgg aac agc ttg ctg tcc gtg gtg cgc gag ggc cgc tcc gta | 6000 |
| Phe Leu Trp Asn Ser Leu Leu Ser Val Val Arg Glu Gly Arg Ser Val |      |
| 1985 1990 1995 2000                                             |      |
| gtg ctc acg tcg cac agc atg gag gag tgc gaa gcg ctc tgc acg cgc | 6048 |
| Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg |      |
| 2005 2010 2015                                                  |      |
| ctg gcc atc atg gtg aac ggg cgg ttc cgc tgt ctg gga agc tct cag | 6096 |
| Leu Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu Gly Ser Ser Gln |      |
| 2020 2025 2030                                                  |      |
| cat ctc aaa ggc agg ttc ggg gct ggc cac aca ctg act cta agg gtc | 6144 |
| His Leu Lys Gly Arg Phe Gly Ala Gly His Thr Leu Thr Leu Arg Val |      |
| 2035 2040 2045                                                  |      |
| cca ccg gac cag cct gag ccg gcg ata gcc ttc atc agg atc aca ttc | 6192 |
| Pro Pro Asp Gln Pro Glu Pro Ala Ile Ala Phe Ile Arg Ile Thr Phe |      |
| 2050 2055 2060                                                  |      |
| cct ggg gct gaa ctc cgg gag gtg cac ggc agc cgt ctg cgc ttc caa | 6240 |
| Pro Gly Ala Glu Leu Arg Glu Val His Gly Ser Arg Leu Arg Phe Gln |      |
| 2065 2070 2075 2080                                             |      |

## FIG. 1K

```

ctg cca ccg ggg ggc aga tgc acc ctg aca cga gtg ttc agg gag ctg      6288
Leu Pro Pro Gly Gly Arg Cys Thr Leu Thr Arg Val Phe Arg Glu Leu
                2085                2090                2095

gct gcc cag ggc agg gcc cac ggt gtg gag gac ttc tct gtg agc cag      6336
Ala Ala Gln Gly Arg Ala His Gly Val Glu Asp Phe Ser Val Ser Gln
                2100                2105                2110

acc act ctg gag gag gtg ttc cta tat ttc tcc aaa gac caa ggg gaa      6384
Thr Thr Leu Glu Glu Val Phe Leu Tyr Phe Ser Lys Asp Gln Gly Glu
                2115                2120                2125

gag gaa gag agc agt cgg cag gag gct gaa gaa gag gag gtt tcc aaa      6432
Glu Glu Glu Ser Ser Arg Gln Glu Ala Glu Glu Glu Glu Val Ser Lys
                2130                2135                2140

cct ggc cgg cag cat ccc aaa cgt gtc agc cga ttc ctg gaa gac ccc      6480
Pro Gly Arg Gln His Pro Lys Arg Val Ser Arg Phe Leu Glu Asp Pro
2145                2150                2155                2160

agc tct gtg gag acc atg atc tga gcatgcctgc cttgggactg agtggcaaag      6534
Ser Ser Val Glu Thr Met Ile
                2165

ctcagacaga ggatctctgt accatacgct ggctcccaga aagccttggg ctctggggga      6594

aataaaaaga aactagaatg agaaaaaaaa aaaaaaaaaa                        6633

```

## FIG. 2A

```

ctcaggggcg gcgcgctccc tgctgtctgc tgggcggagg gaaggcggca agagctgcgg 60
agcccctgga agagcttcca ggaaccctgc gctgtgggat aaaggaatga ggttcagaaa 120
ggggcaggga gttgcccgca gccgcaccgc acgtcttcag cccgaccgtt gtcctgacct 180
ctctgtcccg tccccgccc agtctcacc atg gcc ttc tgg aca cag ctg atg 233
                               Met Ala Phe Trp Thr Gln Leu Met
                               1           5

ctg ctg ctc tgg aag aat ttc atg tat cgc cgg aga cag ccg gtc cag 281
Leu Leu Leu Trp Lys Asn Phe Met Tyr Arg Arg Arg Gln Pro Val Gln
    10           15           20

ctc ctg gtc gaa ttg ctg tgg cct ctc ttc ctc ttc ttc atc ctg gtg 329
Leu Leu Val Glu Leu Leu Trp Pro Leu Phe Leu Phe Phe Ile Leu Val
    25           30           35           40

gct gtt cgc cac tcc cac ccg ccc ctg gag cac cat gaa tgc cac ttc 377
Ala Val Arg His Ser His Pro Pro Leu Glu His His Glu Cys His Phe
                               45           50           55

cca aac aag cca ctg cca tcg gcg ggc acc gtg ccc tgg ctc cag ggt 425
Pro Asn Lys Pro Leu Pro Ser Ala Gly Thr Val Pro Trp Leu Gln Gly
                60           65           70

ctc atc tgt aat gtg aac aac acc tgc ttt ccg cag ctg aca ccg ggc 473
Leu Ile Cys Asn Val Asn Asn Thr Cys Phe Pro Gln Leu Thr Pro Gly
                75           80           85

gag gag ccc ggg cgc ctg agc aac ttc aac gac tcc ctg gtc tcc cgg 521
Glu Glu Pro Gly Arg Leu Ser Asn Phe Asn Asp Ser Leu Val Ser Arg
    90           95           100

ctg cta gcc gat gcc cgc act gtg ctg gga ggg gcc agt gcc cac agg 569
Leu Leu Ala Asp Ala Arg Thr Val Leu Gly Gly Ala Ser Ala His Arg
   105           110           115           120

acg ctg gct ggc cta ggg aag ctg atc gcc acg ctg agg gct gca cgc 617
Thr Leu Ala Gly Leu Gly Lys Leu Ile Ala Thr Leu Arg Ala Ala Arg
                125           130           135

agc acg gcc cag cct caa cca acc aag cag tct cca ctg gaa cca ccc 665
Ser Thr Ala Gln Pro Gln Pro Thr Lys Gln Ser Pro Leu Glu Pro Pro
                140           145           150

atg ctg gat gtc gcg gag ctg ctg acg tca ctg ctg cgc acg gaa tcc 713
Met Leu Asp Val Ala Glu Leu Leu Thr Ser Leu Leu Arg Thr Glu Ser
                155           160           165

ctg ggg ttg gca ctg ggc caa gcc cag gag ccc ttg cac agc ttg ttg 761
Leu Gly Leu Ala Leu Gly Gln Ala Gln Glu Pro Leu His Ser Leu Leu
    170           175           180

```

## FIG. 2B

|                                                                 |      |
|-----------------------------------------------------------------|------|
| gag gcc gct gag gac ctg gcc cag gag ctc ctg gcg ctg cgc agc ctg | 809  |
| Glu Ala Ala Glu Asp Leu Ala Gln Glu Leu Leu Ala Leu Arg Ser Leu |      |
| 185 190 195 200                                                 |      |
| gtg gag ctt cgg gca ctg ctg cag aga ccc cga ggg acc agc ggc ccc | 857  |
| Val Glu Leu Arg Ala Leu Leu Gln Arg Pro Arg Gly Thr Ser Gly Pro |      |
| 205 210 215                                                     |      |
| ctg gag ttg ctg tca gag gcc ctc tgc agt gtc agg gga cct agc agc | 905  |
| Leu Glu Leu Leu Ser Glu Ala Leu Cys Ser Val Arg Gly Pro Ser Ser |      |
| 220 225 230                                                     |      |
| aca gtg ggc ccc tcc ctc aac tgg tac gag gct agt gac ctg atg gag | 953  |
| Thr Val Gly Pro Ser Leu Asn Trp Tyr Glu Ala Ser Asp Leu Met Glu |      |
| 235 240 245                                                     |      |
| ctg gtg ggg cag gag cca gaa tcc gcc ctg cca gac agc agc ctg agc | 1001 |
| Leu Val Gly Gln Glu Pro Glu Ser Ala Leu Pro Asp Ser Ser Leu Ser |      |
| 250 255 260                                                     |      |
| ccc gcc tgc tgc gag ctg att gga gcc ctg gac agc cac ccg ctg tcc | 1049 |
| Pro Ala Cys Ser Glu Leu Ile Gly Ala Leu Asp Ser His Pro Leu Ser |      |
| 265 270 275 280                                                 |      |
| cgc ctg ctc tgg aga cgc ctg aag cct ctg atc ctc ggg aag cta ctc | 1097 |
| Arg Leu Leu Trp Arg Arg Leu Lys Pro Leu Ile Leu Gly Lys Leu Leu |      |
| 285 290 295                                                     |      |
| ttt gca cca gat aca cct ttt acc cgg aag ctc atg gcc cag gtg aac | 1145 |
| Phe Ala Pro Asp Thr Pro Phe Thr Arg Lys Leu Met Ala Gln Val Asn |      |
| 300 305 310                                                     |      |
| cgg acc ttc gag gag ctc acc ctg ctg agg gat gtc cgg gag gtg tgg | 1193 |
| Arg Thr Phe Glu Glu Leu Thr Leu Leu Arg Asp Val Arg Glu Val Trp |      |
| 315 320 325                                                     |      |
| gag atg ctg gga ccc cgg atc ttc acc ttc atg aac gac agt tcc aat | 1241 |
| Glu Met Leu Gly Pro Arg Ile Phe Thr Phe Met Asn Asp Ser Ser Asn |      |
| 330 335 340                                                     |      |
| gtg gcc atg ctg cag cgg ctc ctg cag atg cag gat gaa gga aga agg | 1289 |
| Val Ala Met Leu Gln Arg Leu Leu Gln Met Gln Asp Glu Gly Arg Arg |      |
| 345 350 355 360                                                 |      |
| cag ccc aga cct gga ggc cgg gac cac atg gag gcc ctg cga tcc ttt | 1337 |
| Gln Pro Arg Pro Gly Gly Arg Asp His Met Glu Ala Leu Arg Ser Phe |      |
| 365 370 375                                                     |      |
| ctg gac cct ggg agc ggt ggc tac agc tgg cag gac gca cac gct gat | 1385 |
| Leu Asp Pro Gly Ser Gly Gly Tyr Ser Trp Gln Asp Ala His Ala Asp |      |
| 380 385 390                                                     |      |

## FIG. 2C

|                                                                 |      |
|-----------------------------------------------------------------|------|
| gtg ggg cac ctg gtg ggc acg ctg ggc cga gtg acg gag tgc ctg tcc | 1433 |
| Val Gly His Leu Val Gly Thr Leu Gly Arg Val Thr Glu Cys Leu Ser |      |
| 395 400 405                                                     |      |
| ttg gac aag ctg gag gcg gca ccc tca gag gca gcc ctg gtg tgc cgg | 1481 |
| Leu Asp Lys Leu Glu Ala Ala Pro Ser Glu Ala Ala Leu Val Ser Arg |      |
| 410 415 420                                                     |      |
| gcc ctg caa ctg ctc gcg gaa cat cga ttc tgg gcc ggc gtc gtc ttc | 1529 |
| Ala Leu Gln Leu Leu Ala Glu His Arg Phe Trp Ala Gly Val Val Phe |      |
| 425 430 435 440                                                 |      |
| ttg gga cct gag gac tct tca gac ccc aca gag cac cca acc cca gac | 1577 |
| Leu Gly Pro Glu Asp Ser Ser Asp Pro Thr Glu His Pro Thr Pro Asp |      |
| 445 450 455                                                     |      |
| ctg ggc ccc ggc cac gtg cgc atc aaa atc cgc atg gac att gac gtg | 1625 |
| Leu Gly Pro Gly His Val Arg Ile Lys Ile Arg Met Asp Ile Asp Val |      |
| 460 465 470                                                     |      |
| gtc acg agg acc aat aag atc agg gac agg ttt tgg gac cct ggc cca | 1673 |
| Val Thr Arg Thr Asn Lys Ile Arg Asp Arg Phe Trp Asp Pro Gly Pro |      |
| 475 480 485                                                     |      |
| gcc gcg gac ccc ctg acc gac ctg cgc tac gtg tgg ggc ggc ttc gtg | 1721 |
| Ala Ala Asp Pro Leu Thr Asp Leu Arg Tyr Val Trp Gly Gly Phe Val |      |
| 490 495 500                                                     |      |
| tac ctg caa gac ctg gtg gag cgt gca gcc gtc cgc gtg ctc agc ggc | 1769 |
| Tyr Leu Gln Asp Leu Val Glu Arg Ala Ala Val Arg Val Leu Ser Gly |      |
| 505 510 515 520                                                 |      |
| gcc aac ccc cgg gcc ggc ctc tac ctg cag cag atg ccc tat ccg tgc | 1817 |
| Ala Asn Pro Arg Ala Gly Leu Tyr Leu Gln Gln Met Pro Tyr Pro Cys |      |
| 525 530 535                                                     |      |
| tat gtg gac gac gtg ttc ctg cgt gtg ctg agc cgg tgc ctg ccg ctc | 1865 |
| Tyr Val Asp Asp Val Phe Leu Arg Val Leu Ser Arg Ser Leu Pro Leu |      |
| 540 545 550                                                     |      |
| ttc ctg acg ctg gcc tgg atc tac tcc gtg aca ctg aca gtg aag gcc | 1913 |
| Phe Leu Thr Leu Ala Trp Ile Tyr Ser Val Thr Leu Thr Val Lys Ala |      |
| 555 560 565                                                     |      |
| gtg gtg cgg gag aag gag acg cgg ctg cgg gac acc atg cgc gcc atg | 1961 |
| Val Val Arg Glu Lys Glu Thr Arg Leu Arg Asp Thr Met Arg Ala Met |      |
| 570 575 580                                                     |      |
| ggg ctc agc cgc gcg gtg ctc tgg cta ggc tgg ttc ctc agc tgc ctc | 2009 |
| Gly Leu Ser Arg Ala Val Leu Trp Leu Gly Trp Phe Leu Ser Cys Leu |      |
| 585 590 595 600                                                 |      |

## FIG. 2D

|                                                                 |      |
|-----------------------------------------------------------------|------|
| ggg ccc ttc ctg ctc agc gcc gcg ctg ctg gtt ctg gtg ctc aag ctg | 2057 |
| Gly Pro Phe Leu Leu Ser Ala Ala Leu Leu Val Leu Val Leu Lys Leu |      |
| 605 610 615                                                     |      |
| ggg gac atc ctc ccc tac agc cac ccg ggc gtg gtc ttc ctg ttc ttg | 2105 |
| Gly Asp Ile Leu Pro Tyr Ser His Pro Gly Val Val Phe Leu Phe Leu |      |
| 620 625 630                                                     |      |
| gca gcc ttc gcg gtg gcc acg gtg acc cag agc ttc ctg ctc agc gcc | 2153 |
| Ala Ala Phe Ala Val Ala Thr Val Thr Gln Ser Phe Leu Leu Ser Ala |      |
| 635 640 645                                                     |      |
| ttc ttc tcc cgc gcc aac ctg gct gcg gcc tgc ggc ggc ctg gcc tac | 2201 |
| Phe Phe Ser Arg Ala Asn Leu Ala Ala Ala Cys Gly Gly Leu Ala Tyr |      |
| 650 655 660                                                     |      |
| ttc tcc ctc tac ctg ccc tac gtg ctg tgt gtg gct tgg cgg gac cgg | 2249 |
| Phe Ser Leu Tyr Leu Pro Tyr Val Leu Cys Val Ala Trp Arg Asp Arg |      |
| 665 670 675 680                                                 |      |
| ctg ccc gcg ggt ggc cgc gtg gcc gcg agc ctg ctg tcg ccc gtg gcc | 2297 |
| Leu Pro Ala Gly Gly Arg Val Ala Ala Ser Leu Leu Ser Pro Val Ala |      |
| 685 690 695                                                     |      |
| ttc ggc ttc ggc tgc gag agc ctg gct ctg ctg gag gag cag ggc gag | 2345 |
| Phe Gly Phe Gly Cys Glu Ser Leu Ala Leu Leu Glu Glu Gln Gly Glu |      |
| 700 705 710                                                     |      |
| ggc gcg cag tgg cac aac gtg ggc acc cgg cct acg gca gac gtc ttc | 2393 |
| Gly Ala Gln Trp His Asn Val Gly Thr Arg Pro Thr Ala Asp Val Phe |      |
| 715 720 725                                                     |      |
| agc ctg gcc cag gtc tct ggc ctt ctg ctg ctg gac gcg gcg ctc tac | 2441 |
| Ser Leu Ala Gln Val Ser Gly Leu Leu Leu Leu Asp Ala Ala Leu Tyr |      |
| 730 735 740                                                     |      |
| ggc ctc gcc acc tgg tac ctg gaa gct gtg tgc cca ggc cag tac ggg | 2489 |
| Gly Leu Ala Thr Trp Tyr Leu Glu Ala Val Cys Pro Gly Gln Tyr Gly |      |
| 745 750 755 760                                                 |      |
| atc cct gaa cca tgg aat ttt cct ttt cgg agg agc tac tgg tgc gga | 2537 |
| Ile Pro Glu Pro Trp Asn Phe Pro Phe Arg Arg Ser Tyr Trp Cys Gly |      |
| 765 770 775                                                     |      |
| cct cgg ccc ccc aag agt cca gcc cct tgc ccc acc ccg ctg gac cca | 2585 |
| Pro Arg Pro Pro Lys Ser Pro Ala Pro Cys Pro Thr Pro Leu Asp Pro |      |
| 780 785 790                                                     |      |
| aag gtg ctg gta gaa gag gca ccg ccc ggc ctg agt cct ggc gta tcc | 2633 |
| Lys Val Leu Val Glu Glu Ala Pro Pro Gly Leu Ser Pro Gly Val Ser |      |
| 795 800 805                                                     |      |

## FIG. 2E

|                                                                 |      |
|-----------------------------------------------------------------|------|
| gtt cgc agc ctg gag aag cgc ttt cct gga agc ccg cag cca gcc ctg | 2681 |
| Val Arg Ser Leu Glu Lys Arg Phe Pro Gly Ser Pro Gln Pro Ala Leu |      |
| 810 815 820                                                     |      |
| cgg ggg ctc agc ctg gac ttc tac cag ggc cac atc acc gcc ttc ctg | 2729 |
| Arg Gly Leu Ser Leu Asp Phe Tyr Gln Gly His Ile Thr Ala Phe Leu |      |
| 825 830 835 840                                                 |      |
| ggc cac aac ggg gcc ggc aag acc acc acc ctg tcc atc ttg agt ggc | 2777 |
| Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu Ser Gly |      |
| 845 850 855                                                     |      |
| ctc ttc cca ccc agt ggt ggc tct gcc ttc atc ctg ggc cac gac gtc | 2825 |
| Leu Phe Pro Pro Ser Gly Gly Ser Ala Phe Ile Leu Gly His Asp Val |      |
| 860 865 870                                                     |      |
| cgc tcc agc atg gcc gcc atc cgg ccc cac ctg ggc gtc tgt cct cag | 2873 |
| Arg Ser Ser Met Ala Ala Ile Arg Pro His Leu Gly Val Cys Pro Gln |      |
| 875 880 885                                                     |      |
| tac aac gtg ctg ttt gac atg ctg acc gtg gac gag cac gtc tgg ttc | 2921 |
| Tyr Asn Val Leu Phe Asp Met Leu Thr Val Asp Glu His Val Trp Phe |      |
| 890 895 900                                                     |      |
| tat ggg cgg ctg aag ggt ctg agt gcc gct gta gtg ggc ccc gag cag | 2969 |
| Tyr Gly Arg Leu Lys Gly Leu Ser Ala Ala Val Val Gly Pro Glu Gln |      |
| 905 910 915 920                                                 |      |
| gac cgt ctg ctg cag gat gtg ggg ctg gtc tcc aag cag agt gtg cag | 3017 |
| Asp Arg Leu Leu Gln Asp Val Gly Leu Val Ser Lys Gln Ser Val Gln |      |
| 925 930 935                                                     |      |
| act cgc cac ctc tct ggt ggg atg caa cgg aag ctg tcc gtg gcc att | 3065 |
| Thr Arg His Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val Ala Ile |      |
| 940 945 950                                                     |      |
| gcc ttt gtg ggc ggc tcc caa gtt gtt atc ctg gac gag cct acg gct | 3113 |
| Ala Phe Val Gly Gly Ser Gln Val Val Ile Leu Asp Glu Pro Thr Ala |      |
| 955 960 965                                                     |      |
| ggc gtg gat cct gct tcc cgc cgc ggt att tgg gag ctg ctg ctc aaa | 3161 |
| Gly Val Asp Pro Ala Ser Arg Arg Gly Ile Trp Glu Leu Leu Leu Lys |      |
| 970 975 980                                                     |      |
| tac cga gaa ggt cgc acg ctg atc ctc tcc acc cac cac ctg gat gag | 3209 |
| Tyr Arg Glu Gly Arg Thr Leu Ile Leu Ser Thr His His Leu Asp Glu |      |
| 985 990 995 1000                                                |      |
| gca gag ctg ctg gga gac cgt gtg gct gtg gtg gca ggt ggc cgc ttg | 3257 |
| Ala Glu Leu Leu Gly Asp Arg Val Ala Val Ala Gly Gly Arg Leu     |      |
| 1005 1010 1015                                                  |      |

## FIG. 2F

|                                                                                                                                                           |      |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| tgc tgc tgt ggc tcc cca ctc ttc ctg cgc cgt cac ctg ggc tcc ggc<br>Cys Cys Cys Gly Ser Pro Leu Phe Leu Arg Arg His Leu Gly Ser Gly<br>1020 1025 1030      | 3305 |
| tac tac ctg acg ctg gtg aag gcc cgc ctg ccc ctg acc acc aat gag<br>Tyr Tyr Leu Thr Leu Val Lys Ala Arg Leu Pro Leu Thr Thr Asn Glu<br>1035 1040 1045      | 3353 |
| aag gct gac act gac atg gag ggc agt gtg gac acc agg cag gaa aag<br>Lys Ala Asp Thr Asp Met Glu Gly Ser Val Asp Thr Arg Gln Glu Lys<br>1050 1055 1060      | 3401 |
| aag aat ggc agc cag ggc agc aga gtc ggc act cct cag ctg ctg gcc<br>Lys Asn Gly Ser Gln Gly Ser Arg Val Gly Thr Pro Gln Leu Leu Ala<br>1065 1070 1075 1080 | 3449 |
| ctg gta cag cac tgg gtg ccc ggg gca cgg ctg gtg gag gag ctg cca<br>Leu Val Gln His Trp Val Pro Gly Ala Arg Leu Val Glu Glu Leu Pro<br>1085 1090 1095      | 3497 |
| cac gag ctg gtg ctg gtg ctg ccc tac acg ggt gcc cat gac ggc agc<br>His Glu Leu Val Leu Val Leu Pro Tyr Thr Gly Ala His Asp Gly Ser<br>1100 1105 1110      | 3545 |
| ttc gcc aca ctc ttc cga gag cta gac acg cgg ctg gcg gag ctg agg<br>Phe Ala Thr Leu Phe Arg Glu Leu Asp Thr Arg Leu Ala Glu Leu Arg<br>1115 1120 1125      | 3593 |
| ctc act ggc tac ggg atc tcc gac acc agc ctc gag gag atc ttc ctg<br>Leu Thr Gly Tyr Gly Ile Ser Asp Thr Ser Leu Glu Glu Ile Phe Leu<br>1130 1135 1140      | 3641 |
| aag gtg gtg gag gag tgt gct gcg gac aca gat atg gag gat ggc agc<br>Lys Val Val Glu Glu Cys Ala Ala Asp Thr Asp Met Glu Asp Gly Ser<br>1145 1150 1155 1160 | 3689 |
| tgc ggg cag cac cta tgc aca ggc att gct ggc cta gac gta acc ctg<br>Cys Gly Gln His Leu Cys Thr Gly Ile Ala Gly Leu Asp Val Thr Leu<br>1165 1170 1175      | 3737 |
| cgg ctc aag atg ccg cca cag gag aca gcg ctg gag aac ggg gaa cca<br>Arg Leu Lys Met Pro Pro Gln Glu Thr Ala Leu Glu Asn Gly Glu Pro<br>1180 1185 1190      | 3785 |
| gct ggg tca gcc cca gag act gac cag ggc tct ggg cca gac gcc gtg<br>Ala Gly Ser Ala Pro Glu Thr Asp Gln Gly Ser Gly Pro Asp Ala Val<br>1195 1200 1205      | 3833 |
| ggc cgg gta cag ggc tgg gca ctg acc cgc cag cag ctc cag gcc ctg<br>Gly Arg Val Gln Gly Trp Ala Leu Thr Arg Gln Gln Leu Gln Ala Leu<br>1210 1215 1220      | 3881 |



## FIG. 2H

|                                                                 |      |
|-----------------------------------------------------------------|------|
| ggc cgc tca gtg gag gag ttg tgg gcg ctg ctg agt ccc ctg cct ggc | 4553 |
| Gly Arg Ser Val Glu Glu Leu Trp Ala Leu Leu Ser Pro Leu Pro Gly |      |
| 1435 1440 1445                                                  |      |
| ggg gcc ctc gac cgt gtc ctg aaa aac ctc aca gcc tgg gct cac agc | 4601 |
| Gly Ala Leu Asp Arg Val Leu Lys Asn Leu Thr Ala Trp Ala His Ser |      |
| 1450 1455 1460                                                  |      |
| ctg gat gct cag gac agt ctc aag atc tgg ttc aac aac aaa ggc tgg | 4649 |
| Leu Asp Ala Gln Asp Ser Leu Lys Ile Trp Phe Asn Asn Lys Gly Trp |      |
| 1465 1470 1475 1480                                             |      |
| cac tcc atg gtg gcc ttt gtc aac cga gcc agc aac gca atc ctc cgt | 4697 |
| His Ser Met Val Ala Phe Val Asn Arg Ala Ser Asn Ala Ile Leu Arg |      |
| 1485 1490 1495                                                  |      |
| gct cac ctg ccc cca ggc ccg gcc cgc cac gcc cac agc atc acc aca | 4745 |
| Ala His Leu Pro Pro Gly Pro Ala Arg His Ala His Ser Ile Thr Thr |      |
| 1500 1505 1510                                                  |      |
| ctc aac cac ccc ttg aac ctc acc aag gag cag ctg tct gag gct gca | 4793 |
| Leu Asn His Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ala Ala |      |
| 1515 1520 1525                                                  |      |
| ctg atg gcc tcc tcg gtg gac gtc ctc gtc tcc atc tgt gtg gtc ttt | 4841 |
| Leu Met Ala Ser Ser Val Asp Val Leu Val Ser Ile Cys Val Val Phe |      |
| 1530 1535 1540                                                  |      |
| gcc atg tcc ttt gtc ccg gcc agc ttc act ctt gtc ctc att gag gag | 4889 |
| Ala Met Ser Phe Val Pro Ala Ser Phe Thr Leu Val Leu Ile Glu Glu |      |
| 1545 1550 1555 1560                                             |      |
| cga gtc acc cga gcc aag cac ctg cag ctc atg ggg ggc ctg tcc ccc | 4937 |
| Arg Val Thr Arg Ala Lys His Leu Gln Leu Met Gly Gly Leu Ser Pro |      |
| 1565 1570 1575                                                  |      |
| acc ctc tac tgg ctt ggc aac ttt ctc tgg gac atg tgt aac tac ttg | 4985 |
| Thr Leu Tyr Trp Leu Gly Asn Phe Leu Trp Asp Met Cys Asn Tyr Leu |      |
| 1580 1585 1590                                                  |      |
| gtg cca gca tgc atc gtg gtg ctc atc ttt ctg gcc ttc cag cag agg | 5033 |
| Val Pro Ala Cys Ile Val Val Leu Ile Phe Leu Ala Phe Gln Gln Arg |      |
| 1595 1600 1605                                                  |      |
| gca tat gtg gcc cct gcc aac ctg cct gct ctc ctg ctg ttg cta cta | 5081 |
| Ala Tyr Val Ala Pro Ala Asn Leu Pro Ala Leu Leu Leu Leu Leu     |      |
| 1610 1615 1620                                                  |      |
| ctg tat ggc tgg tcg atc aca ccg ctc atg tac cca gcc tcc ttc ttc | 5129 |
| Leu Tyr Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala Ser Phe Phe |      |
| 1625 1630 1635 1640                                             |      |

## FIG. 2I

```

ttc tcc gtg ccc agc aca gcc tat gtg gtg ctc acc tgc ata aac ctc 5177
Phe Ser Val Pro Ser Thr Ala Tyr Val Val Leu Thr Cys Ile Asn Leu
1645 1650 1655

ttt att ggc atc aat gga agc atg gcc acc ttt gtg ctt gag ctc ttc 5225
Phe Ile Gly Ile Asn Gly Ser Met Ala Thr Phe Val Leu Glu Leu Phe
1660 1665 1670

tct gat cag aag ctg cag gag gtg agc cgg atc ttg aaa cag gtc ttc 5273
Ser Asp Gln Lys Leu Gln Glu Val Ser Arg Ile Leu Lys Gln Val Phe
1675 1680 1685

ctt atc ttc ccc cac ttc tgc ttg ggc cgg ggg ctc att gac atg gtg 5321
Leu Ile Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Met Val
1690 1695 1700

cgg aac cag gcc atg gct gat gcc ttt gag cgc ttg gga gac agg cag 5369
Arg Asn Gln Ala Met Ala Asp Ala Phe Glu Arg Leu Gly Asp Arg Gln
1705 1710 1715 1720

ttc cag tca ccc ctg cgc tgg gag gtg gtc ggc aag aac ctc ttg gcc 5417
Phe Gln Ser Pro Leu Arg Trp Glu Val Val Gly Lys Asn Leu Leu Ala
1725 1730 1735

atg gtg ata cag ggg ccc ctc ttc ctt ctc ttc aca cta ctg ctg cag 5465
Met Val Ile Gln Gly Pro Leu Phe Leu Leu Phe Thr Leu Leu Leu Gln
1740 1745 1750

cac cga agc caa ctc ctg cca cag ccc agg gtg agg tct ctg cca ctc 5513
His Arg Ser Gln Leu Leu Pro Gln Pro Arg Val Arg Ser Leu Pro Leu
1755 1760 1765

ctg gga gag gag gac gag gat gta gcc cgt gaa cgg gag cgg gtg gtc 5561
Leu Gly Glu Glu Asp Glu Asp Val Ala Arg Glu Arg Glu Arg Val Val
1770 1775 1780

caa gga gcc acc cag ggg gat gtg ttg gtg ctg agg aac ttg acc aag 5609
Gln Gly Ala Thr Gln Gly Asp Val Leu Val Leu Arg Asn Leu Thr Lys
1785 1790 1795 1800

gta tac cgt ggg cag agg atg cca gct gtt gac cgc ttg tgc ctg ggg 5657
Val Tyr Arg Gly Gln Arg Met Pro Ala Val Asp Arg Leu Cys Leu Gly
1805 1810 1815

att ccc cct ggt gag tgt ttt ggg ctg ctg ggt gtg aat gga gca ggg 5705
Ile Pro Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly
1820 1825 1830

aag acg tcc acg ttt cgc atg gtg acg ggg gac aca ttg gcc agc agg 5753
Lys Thr Ser Thr Phe Arg Met Val Thr Gly Asp Thr Leu Ala Ser Arg
1835 1840 1845

```

## FIG. 2J

|                                                                 |      |
|-----------------------------------------------------------------|------|
| ggc gag gct gtg ctg gca ggc cac agc gtg gcc cgg gaa ccc agt gct | 5801 |
| Gly Glu Ala Val Leu Ala Gly His Ser Val Ala Arg Glu Pro Ser Ala |      |
| 1850 1855 1860                                                  |      |
| gcg cac ctc agc atg gga tac tgc cct caa tcc gat gcc atc ttt gag | 5849 |
| Ala His Leu Ser Met Gly Tyr Cys Pro Gln Ser Asp Ala Ile Phe Glu |      |
| 1865 1870 1875 1880                                             |      |
| ctg ctg acg ggc cgc gag cac ctg gag ctg ctt gcg cgc ctg cgc ggt | 5897 |
| Leu Leu Thr Gly Arg Glu His Leu Glu Leu Leu Ala Arg Leu Arg Gly |      |
| 1885 1890 1895                                                  |      |
| gtc ccg gag gcc cag gtt gcc cag acc gct ggc tca ggc ctg gcg cgt | 5945 |
| Val Pro Glu Ala Gln Val Ala Gln Thr Ala Gly Ser Gly Leu Ala Arg |      |
| 1900 1905 1910                                                  |      |
| ctg gga ctc tca tgg tac gca gac cgg cct gca ggc acc tac agc gga | 5993 |
| Leu Gly Leu Ser Trp Tyr Ala Asp Arg Pro Ala Gly Thr Tyr Ser Gly |      |
| 1915 1920 1925                                                  |      |
| ggg aac aaa cgc aag ctg gcg acg gcc ctg gcg ctg gtt ggg gac cca | 6041 |
| Gly Asn Lys Arg Lys Leu Ala Thr Ala Leu Ala Leu Val Gly Asp Pro |      |
| 1930 1935 1940                                                  |      |
| gcc gtg gtg ttt ctg gac gag ccg acc aca ggc atg gac ccc agc gcg | 6089 |
| Ala Val Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Ser Ala |      |
| 1945 1950 1955 1960                                             |      |
| cgg cgc ttc ctt tgg aac agc ctt ttg gcc gtg gtg cgg gag ggc cgt | 6137 |
| Arg Arg Phe Leu Trp Asn Ser Leu Leu Ala Val Val Arg Glu Gly Arg |      |
| 1965 1970 1975                                                  |      |
| tca gtg atg ctc acc tcc cat agc atg gag gag tgt gaa gcg ctc tgc | 6185 |
| Ser Val Met Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys |      |
| 1980 1985 1990                                                  |      |
| tcg cgc cta gcc atc atg gtg aat ggg cgg ttc cgc tgc ctg ggc agc | 6233 |
| Ser Arg Leu Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu Gly Ser |      |
| 1995 2000 2005                                                  |      |
| ccg caa cat ctc aag ggc aga ttc gcg gcg ggt cac aca ctg acc ctg | 6281 |
| Pro Gln His Leu Lys Gly Arg Phe Ala Ala Gly His Thr Leu Thr Leu |      |
| 2010 2015 2020                                                  |      |
| cgg gtg ccc gcc gca agg tcc cag ccg gca gcg gcc ttc gtg gcg gcc | 6329 |
| Arg Val Pro Ala Ala Arg Ser Gln Pro Ala Ala Ala Phe Val Ala Ala |      |
| 2025 2030 2035 2040                                             |      |
| gag ttc cct ggg tcg gag ctg cgc gag gca cat gga ggc cgc ctg cgc | 6377 |
| Glu Phe Pro Gly Ser Glu Leu Arg Glu Ala His Gly Gly Arg Leu Arg |      |
| 2045 2050 2055                                                  |      |

## FIG. 2K

```

ttc cag ctg ccg ccg gga ggg cgc tgc gcc ctg gcg cgc gtc ttt gga 6425
Phe Gln Leu Pro Pro Gly Gly Arg Cys Ala Leu Ala Arg Val Phe Gly
                2060                2065                2070

gag ctg gcg gtg cac ggc gca gag cac ggc gtg gag gac ttt tcc gtg 6473
Glu Leu Ala Val His Gly Ala Glu His Gly Val Glu Asp Phe Ser Val
                2075                2080                2085

agc cag acg atg ctg gag gag gta ttc ttg tac ttc tcc aag gac cag 6521
Ser Gln Thr Met Leu Glu Glu Val Phe Leu Tyr Phe Ser Lys Asp Gln
                2090                2095                2100

ggg aag gac gag gac acc gaa gag cag aag gag gca gga gtg gga gtg 6569
Gly Lys Asp Glu Asp Thr Glu Glu Gln Lys Glu Ala Gly Val Gly Val
2105                2110                2115                2120

gac ccc gcg cca ggc ctg cag cac ccc aaa cgc gtc agc cag ttc ctc 6617
Asp Pro Ala Pro Gly Leu Gln His Pro Lys Arg Val Ser Gln Phe Leu
                2125                2130                2135

gat gac cct agc act gcc gag act gtg ctc tga gcctccctcc cctgcggggc 6670
Asp Asp Pro Ser Thr Ala Glu Thr Val Leu
                2140                2145

cgcggggagg ccctgggaat ggcaagggca aggtagagtg cctaggagcc ctggactcag 6730

gctggcagag gggctggtgc cctggagaaa ataaagagaa ggctggagag aagccgtgct 6790

ggtgaaaaaa aaaa 6804

```

## FIG. 3A

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| atg gtc tgc ctg gga act ggc cag agc gct gga ccc cta gtg agt gtt | 48  |
| Met Val Cys Leu Gly Thr Gly Gln Ser Ala Gly Pro Leu Val Ser Val |     |
| 1 5 10 15                                                       |     |
| caa aat cat tgt ccc cct tgt ggt ctt tct ccc cag gaa tcc ctg ggg | 96  |
| Gln Asn His Cys Pro Pro Cys Gly Leu Ser Pro Gln Glu Ser Leu Gly |     |
| 20 25 30                                                        |     |
| ttg gca ctg ggc caa gcc cag gag ccc ttg cac agc ttg ttg gag gcc | 144 |
| Leu Ala Leu Gly Gln Ala Gln Glu Pro Leu His Ser Leu Leu Glu Ala |     |
| 35 40 45                                                        |     |
| gct ggg gac ctg gcc cag gag ctc ctg gcg ctg cgc agc ctg gtg gag | 192 |
| Ala Gly Asp Leu Ala Gln Glu Leu Leu Ala Leu Arg Ser Leu Val Glu |     |
| 50 55 60                                                        |     |
| ctt cgg gca ctg ctg cag aga ccc cga ggg acc agc ggc ccc ctg gag | 240 |
| Leu Arg Ala Leu Leu Gln Arg Pro Arg Gly Thr Ser Gly Pro Leu Glu |     |
| 65 70 75 80                                                     |     |
| ttg ctg tca gag gcc ctc tgc agt gtc agg gga cct agc agc aca gtg | 288 |
| Leu Leu Ser Glu Ala Leu Cys Ser Val Arg Gly Pro Ser Ser Thr Val |     |
| 85 90 95                                                        |     |
| ggc ccc tcc ctc aac tgg tac gag gct agt gac ctg atg gag ctg gtg | 336 |
| Gly Pro Ser Leu Asn Trp Tyr Glu Ala Ser Asp Leu Met Glu Leu Val |     |
| 100 105 110                                                     |     |
| ggg cag gag cca gaa tcc gcc ctg cca gac agc agc ctg agc ccc gcc | 384 |
| Gly Gln Glu Pro Glu Ser Ala Leu Pro Asp Ser Ser Leu Ser Pro Ala |     |
| 115 120 125                                                     |     |
| tgc tgc gag ctg att gga gcc ctg gac agc cac ccg ctg tcc cgc ctg | 432 |
| Cys Ser Glu Leu Ile Gly Ala Leu Asp Ser His Pro Leu Ser Arg Leu |     |
| 130 135 140                                                     |     |
| ctc tgg aga cgc ctg aag cct ctg atc ctc ggg aag cta ctc ttt gca | 480 |
| Leu Trp Arg Arg Leu Lys Pro Leu Ile Leu Gly Lys Leu Leu Phe Ala |     |
| 145 150 155 160                                                 |     |
| cca gat aca cct ttt acc cgg aag ctc atg gcc cag gtg aac cgg acc | 528 |
| Pro Asp Thr Pro Phe Thr Arg Lys Leu Met Ala Gln Val Asn Arg Thr |     |
| 165 170 175                                                     |     |
| ttc gag gag ctc acc ctg ctg agg gat gtc cgg gag gtg tgg gag atg | 576 |
| Phe Glu Glu Leu Thr Leu Leu Arg Asp Val Arg Glu Val Trp Glu Met |     |
| 180 185 190                                                     |     |
| ctg gga ccc cgg atc ttc acc ttc atg aac gac agt tcc aat gtg gcc | 624 |
| Leu Gly Pro Arg Ile Phe Thr Phe Met Asn Asp Ser Ser Asn Val Ala |     |
| 195 200 205                                                     |     |

## FIG. 3B

|                                                                 |      |
|-----------------------------------------------------------------|------|
| atg ctg cag cgg ctc ctg cag atg cag gat gaa gga aga agg cag ccc | 672  |
| Met Leu Gln Arg Leu Leu Gln Met Gln Asp Glu Gly Arg Arg Gln Pro |      |
| 210 215 220                                                     |      |
| aga cct gga ggc cgg gac cac atg gag gcc ctg cga tcc ttt ctg gac | 720  |
| Arg Pro Gly Gly Arg Asp His Met Glu Ala Leu Arg Ser Phe Leu Asp |      |
| 225 230 235 240                                                 |      |
| cct ggg agc ggt ggc tac agc tgg cag gac gca cac gct gat gtg ggg | 768  |
| Pro Gly Ser Gly Gly Tyr Ser Trp Gln Asp Ala His Ala Asp Val Gly |      |
| 245 250 255                                                     |      |
| cac ctg gtg ggc acg ctg ggc cga gtg acg gag tgc ctg tcc ttg gac | 816  |
| His Leu Val Gly Thr Leu Gly Arg Val Thr Glu Cys Leu Ser Leu Asp |      |
| 260 265 270                                                     |      |
| aag ctg gag gcg gca ccc tca gag gca gcc ctg gtg tcg cgg gcc ctg | 864  |
| Lys Leu Glu Ala Ala Pro Ser Glu Ala Ala Leu Val Ser Arg Ala Leu |      |
| 275 280 285                                                     |      |
| caa ctg ctc gcg gaa cat cga ttc tgg gcc ggc gtc gtc ttc ttg gga | 912  |
| Gln Leu Leu Ala Glu His Arg Phe Trp Ala Gly Val Val Phe Leu Gly |      |
| 290 295 300                                                     |      |
| cct gag gac tct tca gac ccc aca gag cac cca acc cca gac ctg ggc | 960  |
| Pro Glu Asp Ser Ser Asp Pro Thr Glu His Pro Thr Pro Asp Leu Gly |      |
| 305 310 315 320                                                 |      |
| ccc ggc cac gtg cgc atc aaa atc cgc atg gac att gac gtg gtc acg | 1008 |
| Pro Gly His Val Arg Ile Lys Ile Arg Met Asp Ile Asp Val Val Thr |      |
| 325 330 335                                                     |      |
| agg acc aat aag atc agg gac agg ttt tgg gac cct ggc cca gcc gcg | 1056 |
| Arg Thr Asn Lys Ile Arg Asp Arg Phe Trp Asp Pro Gly Pro Ala Ala |      |
| 340 345 350                                                     |      |
| gac ccc ctg acc gac ctg cgc tac gtg tgg ggc ggc ttc gtg tac ctg | 1104 |
| Asp Pro Leu Thr Asp Leu Arg Tyr Val Trp Gly Gly Phe Val Tyr Leu |      |
| 355 360 365                                                     |      |
| caa gac ctg gtg gag cgt gca gcc gtc cgc gtg ctc agc ggc gcc aac | 1152 |
| Gln Asp Leu Val Glu Arg Ala Ala Val Arg Val Leu Ser Gly Ala Asn |      |
| 370 375 380                                                     |      |
| ccc cgg gcc ggc ctc tac ctg cag cag atg ccc tat ccg tgc tat gtg | 1200 |
| Pro Arg Ala Gly Leu Tyr Leu Gln Gln Met Pro Tyr Pro Cys Tyr Val |      |
| 385 390 395 400                                                 |      |
| gac gac gtg ttc ctg cgt gtg ctg agc cgg tcg ctg ccg ctc ttc ctg | 1248 |
| Asp Asp Val Phe Leu Arg Val Leu Ser Arg Ser Leu Pro Leu Phe Leu |      |
| 405 410 415                                                     |      |

## FIG. 3C

|                                                                 |      |
|-----------------------------------------------------------------|------|
| acg ctg gcc tgg atc tac tcc gtg aca ctg aca gtg aag gcc gtg gtg | 1296 |
| Thr Leu Ala Trp Ile Tyr Ser Val Thr Leu Thr Val Lys Ala Val Val |      |
| 420 425 430                                                     |      |
| cgg gag aag gag acg cgg ctg cgg gac acc atg cgc gcc atg ggg ctc | 1344 |
| Arg Glu Lys Glu Thr Arg Leu Arg Asp Thr Met Arg Ala Met Gly Leu |      |
| 435 440 445                                                     |      |
| agc cgc gcg gtg ctc tgg cta ggc tgg ttc ctc agc tgc ctc ggg ccc | 1392 |
| Ser Arg Ala Val Leu Trp Leu Gly Trp Phe Leu Ser Cys Leu Gly Pro |      |
| 450 455 460                                                     |      |
| ttc ctg ctc agc gcc gcg ctg ctg gtt ctg gtg ctc aag ctg ggg gac | 1440 |
| Phe Leu Leu Ser Ala Ala Leu Leu Val Leu Val Leu Lys Leu Gly Asp |      |
| 465 470 475 480                                                 |      |
| atc ctc ccc tac agc cac ccg ggc gtg gtc ttc ctg ttc ttg gca gcc | 1488 |
| Ile Leu Pro Tyr Ser His Pro Gly Val Val Phe Leu Phe Leu Ala Ala |      |
| 485 490 495                                                     |      |
| ttc gcg gtg gcc acg gtg acc cag agc ttc ctg ctc agc gcc ttc ttc | 1536 |
| Phe Ala Val Ala Thr Val Thr Gln Ser Phe Leu Leu Ser Ala Phe Phe |      |
| 500 505 510                                                     |      |
| tcc cgc gcc aac ctg gct gcg gcc tgc ggc ggc ctg gcc tac ttc tcc | 1584 |
| Ser Arg Ala Asn Leu Ala Ala Ala Cys Gly Gly Leu Ala Tyr Phe Ser |      |
| 515 520 525                                                     |      |
| ctc tac ctg ccc tac gtg ctg tgt gtg gct tgg cgg gac cgg ctg ccc | 1632 |
| Leu Tyr Leu Pro Tyr Val Leu Cys Val Ala Trp Arg Asp Arg Leu Pro |      |
| 530 535 540                                                     |      |
| gcg ggt ggc cgc gtg gcc gcg agc ctg ctg tcg ccc gtg gcc ttc ggc | 1680 |
| Ala Gly Gly Arg Val Ala Ala Ser Leu Leu Ser Pro Val Ala Phe Gly |      |
| 545 550 555 560                                                 |      |
| ttc ggc tgc gag agc ctg gct ctg ctg gag gag cag ggc gag ggc gcg | 1728 |
| Phe Gly Cys Glu Ser Leu Ala Leu Leu Glu Glu Gln Gly Glu Gly Ala |      |
| 565 570 575                                                     |      |
| cag tgg cac aac gtg ggc acc cgg cct acg gca gac gtc ttc agc ctg | 1776 |
| Gln Trp His Asn Val Gly Thr Arg Pro Thr Ala Asp Val Phe Ser Leu |      |
| 580 585 590                                                     |      |
| gcc cag gtc tct ggc ctt ctg ctg ctg gac gcg gcg ctc tac ggc ctc | 1824 |
| Ala Gln Val Ser Gly Leu Leu Leu Leu Asp Ala Ala Leu Tyr Gly Leu |      |
| 595 600 605                                                     |      |
| gcc acc tgg tac ctg gaa gct gtg tgc cca ggc cag tac ggg atc cct | 1872 |
| Ala Thr Trp Tyr Leu Glu Ala Val Cys Pro Gly Gln Tyr Gly Ile Pro |      |
| 610 615 620                                                     |      |

## FIG. 3D

|                                                                 |      |
|-----------------------------------------------------------------|------|
| gaa cca tgg aat ttt cct ttt cgg agg agc tac tgg tgc gga cct cgg | 1920 |
| Glu Pro Trp Asn Phe Pro Phe Arg Arg Ser Tyr Trp Cys Gly Pro Arg |      |
| 625 630 635 640                                                 |      |
| ccc ccc aag agt cca gcc cct tgc ccc acc ccg ctg gac cca aag gtg | 1968 |
| Pro Pro Lys Ser Pro Ala Pro Cys Pro Thr Pro Leu Asp Pro Lys Val |      |
| 645 650 655                                                     |      |
| ctg gta gaa gag gca ccg ccc ggc ctg agt cct ggc gta tcc gtt cgc | 2016 |
| Leu Val Glu Glu Ala Pro Pro Gly Leu Ser Pro Gly Val Ser Val Arg |      |
| 660 665 670                                                     |      |
| agc ctg gag aag cgc ttt cct gga agc ccg cag cca gcc ctg cgg ggg | 2064 |
| Ser Leu Glu Lys Arg Phe Pro Gly Ser Pro Gln Pro Ala Leu Arg Gly |      |
| 675 680 685                                                     |      |
| ctc agc ctg gac ttc tac cag ggc cac atc acc gcc ttc ctg ggc cac | 2112 |
| Leu Ser Leu Asp Phe Tyr Gln Gly His Ile Thr Ala Phe Leu Gly His |      |
| 690 695 700                                                     |      |
| aac ggg gcc ggc aag acc acc acc ctg tcc atc ttg agt ggc ctc ttc | 2160 |
| Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu Ser Gly Leu Phe |      |
| 705 710 715 720                                                 |      |
| cca ccc agt ggt ggc tct gcc ttc atc ctg ggc cac gac gtc cgc tcc | 2208 |
| Pro Pro Ser Gly Gly Ser Ala Phe Ile Leu Gly His Asp Val Arg Ser |      |
| 725 730 735                                                     |      |
| agc atg gcc gcc atc cgg ccc cac ctg ggc gtc tgt cct cag tac aac | 2256 |
| Ser Met Ala Ala Ile Arg Pro His Leu Gly Val Cys Pro Gln Tyr Asn |      |
| 740 745 750                                                     |      |
| gtg ctg ttt gac atg ctg acc gtg gac gag cac gtc tgg ttc tat ggg | 2304 |
| Val Leu Phe Asp Met Leu Thr Val Asp Glu His Val Trp Phe Tyr Gly |      |
| 755 760 765                                                     |      |
| cgg ctg aag ggt ctg agt gcc gct gta gtg ggc ccc gag cag gac cgt | 2352 |
| Arg Leu Lys Gly Leu Ser Ala Ala Val Val Gly Pro Glu Gln Asp Arg |      |
| 770 775 780                                                     |      |
| ctg ctg cag gat gtg ggg ctg gtc tcc aag cag agt gtg cag act cgc | 2400 |
| Leu Leu Gln Asp Val Gly Leu Val Ser Lys Gln Ser Val Gln Thr Arg |      |
| 785 790 795 800                                                 |      |
| cac ctc tct ggt ggg atg caa cgg aag ctg tcc gtg gcc att gcc ttt | 2448 |
| His Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val Ala Ile Ala Phe |      |
| 805 810 815                                                     |      |
| gtg ggc ggc tcc caa gtt gtt atc ctg gac gag cct acg gct ggc gtg | 2496 |
| Val Gly Gly Ser Gln Val Val Ile Leu Asp Glu Pro Thr Ala Gly Val |      |
| 820 825 830                                                     |      |

## FIG. 3E

|                                                                 |      |
|-----------------------------------------------------------------|------|
| gat cct gct tcc cgc cgc ggt att tgg gag ctg ctg ctc aaa tac cga | 2544 |
| Asp Pro Ala Ser Arg Arg Gly Ile Trp Glu Leu Leu Leu Lys Tyr Arg |      |
| 835 840 845                                                     |      |
| gaa ggt cgc acg ctg atc ctc tcc acc cac cac ctg gat gag gca gag | 2592 |
| Glu Gly Arg Thr Leu Ile Leu Ser Thr His His Leu Asp Glu Ala Glu |      |
| 850 855 860                                                     |      |
| ctg ctg gga gac cgt gtg gct gtg gtg gca ggt ggc cgc ttg tgc tgc | 2640 |
| Leu Leu Gly Asp Arg Val Ala Val Val Ala Gly Gly Arg Leu Cys Cys |      |
| 865 870 875 880                                                 |      |
| tgt ggc tcc cca ctc ttc ctg cgc cgt cac ctg ggc tcc ggc tac tac | 2688 |
| Cys Gly Ser Pro Leu Phe Leu Arg Arg His Leu Gly Ser Gly Tyr Tyr |      |
| 885 890 895                                                     |      |
| ctg acg ctg gtg aag gcc cgc ctg ccc ctg acc acc aat gag aag gct | 2736 |
| Leu Thr Leu Val Lys Ala Arg Leu Pro Leu Thr Thr Asn Glu Lys Ala |      |
| 900 905 910                                                     |      |
| gac act gac atg gag ggc agt gtg gac acc agg cag gaa aag aag aat | 2784 |
| Asp Thr Asp Met Glu Gly Ser Val Asp Thr Arg Gln Glu Lys Lys Asn |      |
| 915 920 925                                                     |      |
| ggc agc cag ggc agc aga gtc ggc act cct cag ctg ctg gcc ctg gta | 2832 |
| Gly Ser Gln Gly Ser Arg Val Gly Thr Pro Gln Leu Leu Ala Leu Val |      |
| 930 935 940                                                     |      |
| cag cac tgg gtg ccc ggg gca cgg ctg gtg gag gag ctg cca cac gag | 2880 |
| Gln His Trp Val Pro Gly Ala Arg Leu Val Glu Glu Leu Pro His Glu |      |
| 945 950 955 960                                                 |      |
| ctg gtg ctg gtg ctg ccc tac acg ggt gcc cat gac ggc agc ttc gcc | 2928 |
| Leu Val Leu Val Leu Pro Tyr Thr Gly Ala His Asp Gly Ser Phe Ala |      |
| 965 970 975                                                     |      |
| aca ctc ttc cga gag cta gac acg cgg ctg gcg gag ctg agg ctc act | 2976 |
| Thr Leu Phe Arg Glu Leu Asp Thr Arg Leu Ala Glu Leu Arg Leu Thr |      |
| 980 985 990                                                     |      |
| ggc tac ggg atc tcc gac acc agc ctc gag gag atc ttc ctg aag gtg | 3024 |
| Gly Tyr Gly Ile Ser Asp Thr Ser Leu Glu Glu Ile Phe Leu Lys Val |      |
| 995 1000 1005                                                   |      |
| gtg gag gag tgt gct gcg gac aca gat atg gag gat ggc agc tgc ggg | 3072 |
| Val Glu Glu Cys Ala Ala Asp Thr Asp Met Glu Asp Gly Ser Cys Gly |      |
| 1010 1015 1020                                                  |      |
| cag cac cta tgc aca ggc att gct ggc cta gac gta acc ctg cgg ctc | 3120 |
| Gln His Leu Cys Thr Gly Ile Ala Gly Leu Asp Val Thr Leu Arg Leu |      |
| 1025 1030 1035 1040                                             |      |

## FIG. 3F

|                                                                 |      |
|-----------------------------------------------------------------|------|
| aag atg ccg cca cag gag aca gcg ctg gag aac ggg gaa cca gct ggg | 3168 |
| Lys Met Pro Pro Gln Glu Thr Ala Leu Glu Asn Gly Glu Pro Ala Gly |      |
| 1045 1050 1055                                                  |      |
| tca gcc cca gag act gac cag ggc tct ggg cca gac gcc gtg ggc cgg | 3216 |
| Ser Ala Pro Glu Thr Asp Gln Gly Ser Gly Pro Asp Ala Val Gly Arg |      |
| 1060 1065 1070                                                  |      |
| gta cag ggc tgg gca ctg acc cgc cag cag ctc cag gcc ctg ctt ctc | 3264 |
| Val Gln Gly Trp Ala Leu Thr Arg Gln Gln Leu Gln Ala Leu Leu Leu |      |
| 1075 1080 1085                                                  |      |
| aag cgc ttt ctg ctt gcc cgc cgc agc cgc cgc ggc ctg ttc gcc cag | 3312 |
| Lys Arg Phe Leu Leu Ala Arg Arg Ser Arg Arg Gly Leu Phe Ala Gln |      |
| 1090 1095 1100                                                  |      |
| atc gtg ctg cct gcc ctc ttt gtg ggc ctg gcc ctc gtg ttc agc ctc | 3360 |
| Ile Val Leu Pro Ala Leu Phe Val Gly Leu Ala Leu Val Phe Ser Leu |      |
| 1105 1110 1115 1120                                             |      |
| atc gtg cct cct ttc ggg cac tac ccg gct ctg cgg ctc agt ccc acc | 3408 |
| Ile Val Pro Pro Phe Gly His Tyr Pro Ala Leu Arg Leu Ser Pro Thr |      |
| 1125 1130 1135                                                  |      |
| atg tac ggt gct cag gtg tcc ttc ttc agt gag gac gcc cca ggg gac | 3456 |
| Met Tyr Gly Ala Gln Val Ser Phe Phe Ser Glu Asp Ala Pro Gly Asp |      |
| 1140 1145 1150                                                  |      |
| cct gga cgt gcc cgg ctg ctc gag gcg ctg ctg cag gag gca gga ctg | 3504 |
| Pro Gly Arg Ala Arg Leu Leu Glu Ala Leu Leu Gln Glu Ala Gly Leu |      |
| 1155 1160 1165                                                  |      |
| gag gag ccc cca gtg cag cat agc tcc cac agg ttc tcg gca cca gaa | 3552 |
| Glu Glu Pro Pro Val Gln His Ser Ser His Arg Phe Ser Ala Pro Glu |      |
| 1170 1175 1180                                                  |      |
| gtt cct gct gaa gtg gcc aag gtc ttg gcc agt ggc aac tgg acc cca | 3600 |
| Val Pro Ala Glu Val Ala Lys Val Leu Ala Ser Gly Asn Trp Thr Pro |      |
| 1185 1190 1195 1200                                             |      |
| gag tct cca tcc cca gcc tgc cag tgt agc cag ccc ggt gcc cgg cgc | 3648 |
| Glu Ser Pro Ser Pro Ala Cys Gln Cys Ser Gln Pro Gly Ala Arg Arg |      |
| 1205 1210 1215                                                  |      |
| ctg ctg ccc gac tgc ccg gct gca gct ggt ggt ccc cct ccg ccc cag | 3696 |
| Leu Leu Pro Asp Cys Pro Ala Ala Ala Gly Gly Pro Pro Pro Gln     |      |
| 1220 1225 1230                                                  |      |
| gca gtg acc ggc tct ggg gaa gtg gtt cag aac ctg aca ggc cgg aac | 3744 |
| Ala Val Thr Gly Ser Gly Glu Val Val Gln Asn Leu Thr Gly Arg Asn |      |
| 1235 1240 1245                                                  |      |

## FIG. 3G

|                                                                 |      |
|-----------------------------------------------------------------|------|
| ctg tct gac ttc ctg gtc aag acc tac ccg cgc ctg gtg cgc cag ggc | 3792 |
| Leu Ser Asp Phe Leu Val Lys Thr Tyr Pro Arg Leu Val Arg Gln Gly |      |
| 1250 1255 1260                                                  |      |
| ctg aag act aag aag tgg gtg aat gag gtc agg tac gga ggc ttc tcg | 3840 |
| Leu Lys Thr Lys Lys Trp Val Asn Glu Val Arg Tyr Gly Gly Phe Ser |      |
| 1265 1270 1275 1280                                             |      |
| ctg ggg ggc cga gac cca ggc ctg ccc tcg ggc caa gag ttg ggc cgc | 3888 |
| Leu Gly Gly Arg Asp Pro Gly Leu Pro Ser Gly Gln Glu Leu Gly Arg |      |
| 1285 1290 1295                                                  |      |
| tca gtg gag gag ttg tgg gcg ctg ctg agt ccc ctg cct ggc ggg gcc | 3936 |
| Ser Val Glu Glu Leu Trp Ala Leu Leu Ser Pro Leu Pro Gly Gly Ala |      |
| 1300 1305 1310                                                  |      |
| ctc gac cgt gtc ctg aaa aac ctc aca gcc tgg gct cac agc ctg gat | 3984 |
| Leu Asp Arg Val Leu Lys Asn Leu Thr Ala Trp Ala His Ser Leu Asp |      |
| 1315 1320 1325                                                  |      |
| gct cag gac agt ctc aag atc tgg ttc aac aac aaa ggc tgg cac tcc | 4032 |
| Ala Gln Asp Ser Leu Lys Ile Trp Phe Asn Asn Lys Gly Trp His Ser |      |
| 1330 1335 1340                                                  |      |
| atg gtg gcc ttt gtc aac cga gcc agc aac gca atc ctc cgt gct cac | 4080 |
| Met Val Ala Phe Val Asn Arg Ala Ser Asn Ala Ile Leu Arg Ala His |      |
| 1345 1350 1355 1360                                             |      |
| ctg ccc cca ggc ccg gcc cgc cac gcc cac agc atc acc aca ctc aac | 4128 |
| Leu Pro Pro Gly Pro Ala Arg His Ala His Ser Ile Thr Thr Leu Asn |      |
| 1365 1370 1375                                                  |      |
| cac ccc ttg aac ctc acc aag gag cag ctg tct gag gct gca ctg atg | 4176 |
| His Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ala Ala Leu Met |      |
| 1380 1385 1390                                                  |      |
| gcc tcc tcg gtg gac gtc ctc gtc tcc atc tgt gtg gtc ttt gcc atg | 4224 |
| Ala Ser Ser Val Asp Val Leu Val Ser Ile Cys Val Val Phe Ala Met |      |
| 1395 1400 1405                                                  |      |
| tcc ttt gtc ccg gcc agc ttc act ctt gtc ctc att gag gag cga gtc | 4272 |
| Ser Phe Val Pro Ala Ser Phe Thr Leu Val Leu Ile Glu Glu Arg Val |      |
| 1410 1415 1420                                                  |      |
| acc cga gcc aag cac ctg cag ctc atg ggg ggc ctg tcc ccc acc ctc | 4320 |
| Thr Arg Ala Lys His Leu Gln Leu Met Gly Gly Leu Ser Pro Thr Leu |      |
| 1425 1430 1435 1440                                             |      |
| tac tgg ctt ggc aac ttt ctc tgg gac atg tgt aac tac ttg gtg cca | 4368 |
| Tyr Trp Leu Gly Asn Phe Leu Trp Asp Met Cys Asn Tyr Leu Val Pro |      |
| 1445 1450 1455                                                  |      |

## FIG. 3H

|                                                                 |      |
|-----------------------------------------------------------------|------|
| gca tgc atc gtg gtg ctc atc ttt ctg gcc ttc cag cag agg gca tat | 4416 |
| Ala Cys Ile Val Val Leu Ile Phe Leu Ala Phe Gln Gln Arg Ala Tyr |      |
| 1460 1465 1470                                                  |      |
| gtg gcc cct gcc aac ctg cct gct ctc ctg ctg ttg cta cta ctg tat | 4464 |
| Val Ala Pro Ala Asn Leu Pro Ala Leu Leu Leu Leu Leu Leu Tyr     |      |
| 1475 1480 1485                                                  |      |
| ggg aga cag gca gtt cca gtc acc cct gcg ctg gga ggt ggt cgg caa | 4512 |
| Gly Arg Gln Ala Val Pro Val Thr Pro Ala Leu Gly Gly Gly Arg Gln |      |
| 1490 1495 1500                                                  |      |
| gaa cct ctt ggc cat ggt gat aca ggg gcc cct ctt cct tct ctt cac | 4560 |
| Glu Pro Leu Gly His Gly Asp Thr Gly Ala Pro Leu Pro Ser Leu His |      |
| 1505 1510 1515 1520                                             |      |
| act act gct gca gca ccg aag cca act cct gcc aca gcc cag ggt gag | 4608 |
| Thr Thr Ala Ala Ala Pro Lys Pro Thr Pro Ala Thr Ala Gln Gly Glu |      |
| 1525 1530 1535                                                  |      |
| gtc tct gcc act cct ggg aga gga gga cga gga tgt agc ccg tga     | 4653 |
| Val Ser Ala Thr Pro Gly Arg Gly Gly Arg Gly Cys Ser Pro         |      |
| 1540 1545 1550                                                  |      |

## FIG. 4A

|        |            |             |                                   |
|--------|------------|-------------|-----------------------------------|
|        | 1          |             | 50                                |
| muABC1 | .....      | .....       | .....                             |
| muABCR | MGFLRQIQLL | LWKNWTLRKR  | QKIRFVVELV WPLSLFLVLI WLRNANPLYS  |
| muABCL | MAFCTQLMLL | LWKNYTYRRR  | QPIQLLVELL WPLFLFFILV AVRHSHPPLE  |
|        | 51         |             | 100                               |
| muABC1 | .....      | MPSAGTLPWV  | QGIICNANNP CFPYPTPGEA PGVVGNFNKS  |
| muABCR | QHECHFNPKA | MPSAGLLPWL  | QGIFCNMNNP CFQNPTPGES PGTVSNNYNS  |
| muABCL | HHECHFNPKP | LPSAGTVPWL  | QGLVCNVNNS CFQHPTPGEK PGVLSNFKDS  |
|        | 101        |             | 150                               |
| muABC1 | IVSRLFSDAQ | RLLLYSQRDT  | SIKDMHKVLR MLRQIKHPN. ....SN      |
| muABCR | ILARVYRDFQ | ELFMDTPEVQ  | HLGQVWAE LR TLSQFMDTLR THPERFAGRG |
| muABCL | LISRLADTR  | TVLGGHSIQD  | MLDALGKLIP VLRVGGGA. ....         |
|        | 151        |             | 200                               |
| muABC1 | LKLQDFLVDN | ETFSGFLQHN  | LSLPRSTVDS LLQXNVGLQK VFLQGYQLHL  |
| muABCR | LQIRDILKDE | EALTFLFLMRN | IGLSDSVAHL LVNSQVRVEQ FAYGVPDLEL  |
| muABCL | .....      | .....       | .....                             |
|        | 201        |             | 250                               |
| muABC1 | ASLCNGSKLE | EIIQLG....  | ..DAEVSALC GLPRKKLDAA ERVLRYNMDI  |
| muABCR | TDIACSEALL | QRFIIFSQRR  | GAQTVRDALC PLSQVTLQWI EDTLYADVDF  |
| muABCL | .....      | .....       | ..RPQESDQP TSQGSVTKLL EKILQRASLD  |
|        | 251        |             | 300                               |
| muABC1 | LKPVVTKLNS | TSHLPTQHLA  | EATTVLLDSL GGLAQELFST KSWSDMRQEV  |
| muABCR | FKLFHVLPTL | LDSSSQGINL  | RFWGGILSDL SPRMQKFIHR PSVQDLLWVS  |
| muABCL | PVLGQAQDSM | RKFSDAIRDL  | AQELLTLPSL MELRALLR.. R.....      |
|        | 301        |             | 350                               |
| muABC1 | MFLTNNVSSS | SSTQIYQAVS  | RIVCGHPEGG GLKIKSLNWX EDNNYKALFG  |
| muABCR | RPLLQNGGPE | TFTQLMSILS  | DLICGYPEGG GSRVFSFNWX EDNNYKAFLG  |
| muABCL | .....PR    | GSAGSLELVS  | EALCSTKGPS SPGGLSLNWX EANQLNEFMG  |
|        | 351        |             | 400                               |
| muABC1 | GNNTEEDVDT | FYDNSTTPYC  | NDLMKNLESS PLSRIIWKAL KPLLVGKILY  |
| muABCR | IDSTRKDPAY | SYDKRTTSEC  | NSLIQSLESN PLTKIAWRAA KPLLMGKILF  |
| muABCL | ....PEVAPA | LPDNLSPAC   | SEFVGTLDDH PVSRLWRRRL KPLLKGKILF  |
|        | 401        |             | 450                               |
| muABC1 | TPDTPATRQV | MAEVNKTQFE  | LAVFHDLEGM WEELSPQIWT FMENSQEMDL  |
| muABCR | TPDSPAARRI | MKNANSTFEE  | LDRVRKLVKA WEEVGPQIWX FFEKSTQMTV  |
| muABCL | APDTNFTRKL | MAQVNQTFEE  | LALLRDLHEL WGVLPQIFN FMNDSTNVAM   |
|        | 451        |             | 500                               |
| muABC1 | VRTLLDSRGN | DQFWEQKLDG  | LDWTAQDIMA FLAKNPEDVQ SPNGSVYTWR  |
| muABCR | IRDTLQHPTV | KDFINRQLGE  | EGITTEAVLN FFSNGPQEKQ ADDMTSFDWR  |
| muABCL | LQRLLDVGGT | GQRQQTTPRAQ | .....K.. KL.EAIKDFL DPSRGGYSWR    |

## FIG. 4B

|        |            |            |                                   |
|--------|------------|------------|-----------------------------------|
|        | 501        |            | 550                               |
| muABC1 | EAUNETNQAI | QTISRMECEV | NLNKLEPIPT EVRLINKSME LLDERKFWAG  |
| muABCR | DIFNITDRFL | RLANQYLECL | VLDKFESYDD EVQLTQRALS LLEENRFWAG  |
| muABCL | EAHADMGRLA | GILGQMMCEV | SLDKLEAVPS EEALVSRALE LLGERRLWAG  |
|        | 551        |            | 600                               |
| muABC1 | IVFTG..... | ...ITPDSVE | LPHHVKYKIR MDIDNVERTN KIKDGYWDPG  |
| muABCR | VVFPG..... | ...MYPWASS | LPPHVKYKIR MDIDVVEKTN KIKDRYWDSG  |
| muABCL | IVFLSPEHPL | DPSELSSPAL | SPGHLRFKIR MDIDDVTRTN KIRDKFWDPG  |
|        | 601        |            | 650                               |
| muABC1 | PRADPFEDMR | YVWGGFAYLQ | DVVEQAIIRV LTGSEKKTGV YVQQMPYPYCY |
| muABCR | PRADPVEDFR | YIWGGFAYLQ | DMVEQGIVKS QMQAEPPIGV YLQQMPYPYCF |
| muABCL | PSADPFMDLR | YVWGGFVYLQ | DLLEQAARV LGGGNSRTGL YLQQMPHPYCY  |
|        | 651        |            | 700                               |
| muABC1 | VDDIFLRVMS | RSMPLFMTLA | WIYSVAVIK SIVYEKEARL KETMRIMGLD   |
| muABCR | VDDSFMIILN | RCFPIFMVLA | WIYSVSMTVK GIVLEKELRL KETLKNQGV   |
| muABCL | VDDVFLRVLS | RSLPLFLTLA | WIYSVALTVK AVVREKETRL RETMRAMGLS  |
|        | 701        |            | 750                               |
| muABC1 | NGILWFSWFV | SSLIPLLVSA | GLLVVILKLG NLLPYSDPSV VFVFLSVFAM  |
| muABCR | NAVIWCTWFL | DSFSIMALSI | FLLTLFIMHG RILHYSDFPI LFLFLLAFAF  |
| muABCL | RAVLWLGWFL | SCLGPFLVSA | ALLVLVLKLG NILPYSHPVV IFLFLAAFAV  |
|        | 751        |            | 800                               |
| muABC1 | VTILQCFLIS | TLFSRANLAA | ACGGIIYFTL YLPYVLCVAW QDYVGFSIKI  |
| muABCR | ATIMQSFLLS | TLFSKASLAA | ACSGVIYFTL YLPHVLCFAW QDRMTADLKT  |
| muABCL | ATVAQSFLLS | AFFSRANLAA | ACGGLAYFAL YLPYVLCVAW RERLHLGGLL  |
|        | 801        |            | 850                               |
| muABC1 | FASLLSPVAF | GFGCEYFALF | EEQGIGVQWD NLFESPVEED GFNLTTAVSM  |
| muABCR | TVSLLSSVAF | GFGTEYLVRF | EEQGLGLQWS NIGKSPLEGD EFSLLSMKM   |
| muABCL | AASLLSPVAF | GFGCESLALL | EEQGDGAQWH NLGTG.PAED VFSLAQVSAF  |
|        | 851        |            | 900                               |
| muABC1 | MLFDTFLYGV | MTWYIEAVFP | GQYGIPRPWY FPCTKSYWFG .....       |
| muABCR | MLLDAALYGL | LAWYLDQVFP | GDYGTPLPWY FLLQESYWL GEGCSTREER   |
| muABCL | LLLDAVIYGL | ALWYLEAVCP | GQYGIPEPWN FPFRRSYWCG .....       |
|        | 901        |            | 950                               |
| muABC1 | .....EEIDE | KSHPGSSQKG | VSEICMEEEP THLRLGVSIQ NLVKVYRDGM  |
| muABCR | ALEKTEPLTE | EMEDPEHPEG | MNDSFFEREL PGLVPGVCVK NLVKVFEP    |
| muABCL | .....PGP.P | KSSVLAPAPQ | DPKVLVEEPP LGLVPGVSIR GLKKHFRGCP  |
|        | 951        |            | 1000                              |
| muABC1 | KVAVDGLALN | FYEGQITSFL | GHNGAGKTTT MSILTGLFPP TSGTAYILGK  |
| muABCR | RPAVDRLNIT | FYENQITAF  | LGHNGAGKTTT LSILTGLLPP TSGTVLIGGK |
| muABCL | QPALQGLNLD | FYEGHITAF  | LGHNGAGKTTT LSILSGLFPP SSGSASILGH |

## FIG. 4C

|        |                                                          |  |      |
|--------|----------------------------------------------------------|--|------|
|        | 1001                                                     |  | 1050 |
| muABC1 | DIRSEMSSIR QNLGVCPQHN VLFDMLTVEE HIWFIYARLKG LSEKHVKAEM  |  |      |
| muABCR | DIETNLDVVR QSLGMCPQHN ILFHHLTVAE HILFYAQLKG RSWEEAQLEM   |  |      |
| muABCL | DVQTNMAAIR PHLGICPQYN VLFDMLTVEE HVWFIYGRLLKG VSAAAMGPER |  |      |
|        | 1051                                                     |  | 1100 |
| muABC1 | EQMALDVGLP PSKLKSKTSQ LSGGMQRKLS VALAFVGGSK VVILDEPTAG   |  |      |
| muABCR | EAMLEDTGLH HK.RNEEAQD LSGGMQRKLS VAIAFVGDSK VVVLDEPTSG   |  |      |
| muABCL | ERLIRDVGLT LK.RDTQTRH LSGGMQRKLS VAIAFVGDSK VVIMDEPTAG   |  |      |
|        | 1101                                                     |  | 1150 |
| muABC1 | VDPYSRRGIW ELLLKYRQGR TIILSTHHMD EADILGDRIA IISHGKLCCV   |  |      |
| muABCR | VDPYSRRSIW DLLLKYRSGR TIIMSTHHMD EADLLGDRIA IISQGRLYCS   |  |      |
| muABCL | VDPASRRGIW ELLLKYREGR TLILSTHHLD EAELLGDRVA MVAGGSLCCC   |  |      |
|        | 1151                                                     |  | 1200 |
| muABC1 | GSSLFLKNQL GTGYLTLVK KDVESSLSSC RNSSSTVSCL KKEDSVSQSS    |  |      |
| muABCR | GTPLFLKNCF GTGFYTLVLR KMKNIQSQRG .GCEGVCSCS SKGFSTR...   |  |      |
| muABCL | GSPLFLRRHL GCGYTLTLVK SSQSLVTHDA KGDSEDPRE KKSDBGNGRTS   |  |      |
|        | 1201                                                     |  | 1250 |
| muABC1 | .....S DAGLGSDHES DTLTIDVSAI SNLIRKHVSE ARLVEDIGHE       |  |      |
| muABCR | .....C PTRVDEITEE QVLDGDVQEL MDLVYHHVPE AKLVEICIGQE      |  |      |
| muABCL | DTAFTRGTSD KSNQAPAPGA VPITPSTARI LELVQQHVPK AQLVEDLPHE   |  |      |
|        | 1251                                                     |  | 1300 |
| muABC1 | LTIVLPYEAA KEGAFVELFH EIDDRLSDLG ISSYGISETT LEEIFLKVAE   |  |      |
| muABCR | LIFLLPNKNF KQRAYASLFR ELEETLADLG LSSFGISDTP LEEIFLKVTE   |  |      |
| muABCL | LLLVLPHYAGA LDGSFAMVFQ ELDQQLLELLG LTGYGISDTN LEEIFLKVVE |  |      |
|        | 1301                                                     |  | 1350 |
| muABC1 | ESGVDAETSD GTLPARRNR AFGDKQSCLH PFTEDDAVDP NDSIDDPESR    |  |      |
| muABCR | DAGAGSMFVG GAQQKREQAG LRHPCSAPE KLRQYAQAPH TCSPGQVDP     |  |      |
| muABCL | DA...HREGG DSRPQLHLR. .TCTPQPPTG PEASVLENGE LAK...LV..L  |  |      |
|        | 1351                                                     |  | 1400 |
| muABC1 | ETDLLSGMDG KGSYQLKGWK LTQQQFVALL WKRLLIARRS RKGFFAQIVL   |  |      |
| muABCR | KGQPSPEPED PGVPFNTGAR LILQHVQALL VKRFHHTIRS RKDFVAQIVL   |  |      |
| muABCL | D.PQAPQGLA PNAAQVQGW LTCQQLRALL HKRFLIARRS RRGLFAQVVL    |  |      |
|        | 1401                                                     |  | 1450 |
| muABC1 | PAVFVCIALV FSLIVPPFGK YPSLELQPWM YNEQYTFVSN DAPEDMGTQE   |  |      |
| muABCR | PATFVFLALM LSIIVPPFGE FPALTLPWM YGHQYTFFSM DEPNNEHLEV    |  |      |
| muABCL | PALFVGLALF FSLIVPPFGQ YPPLQLSPAM YGPQVSFFSE DAPGDPNRMK   |  |      |
|        | 1451                                                     |  | 1500 |
| muABC1 | LLNALTKDPG FGTRCMENP IPDTPCLAG. EEDWTISPVP QSIVDLFQNG    |  |      |
| muABCR | LADVLLNRPG FGNRCLKEEW LPEYPCIN.. ATSWKTPSVS PNITHLFQKQ   |  |      |
| muABCL | LLEALLGEAG LQEPSMQDKD ARGSECTHSL ACYFTVPEVP PDVASILASG   |  |      |

## FIG. 4D

|        |             |            |                                     |
|--------|-------------|------------|-------------------------------------|
|        | 1501        |            | 1550                                |
| muABC1 | NWTMKNPSPA  | CQCSSDKIKK | MLPVCPPGAG GLPPPQRKQK TADILQNLTG    |
| muABCR | KWTAHPSPS   | CKCSTREKLT | MLPECPEGAG GLPPPQRTQR STEVLQDLTN    |
| muABCL | NWTPESPSPA  | CQCSQPGARR | LLPDCPAGAG GPPPPQAVAG LGEVVQNLTG    |
|        | 1551        |            | 1600                                |
| muABC1 | RNISDYL VKT | YVQIIAKSLK | NKIWVNEFRY GGFSLGVSNS QALPPSHEVN    |
| muABCR | RNISDYL VKT | YPALIRSSLK | SKFWVNEQRY GGISIGGKLP AIPISGEALV    |
| muABCL | RNVSDFL VKT | YPSLVRRGLK | TKKWVDEVRY GGFSLGGRDP .DLPTGHEVV    |
|        | 1601        |            | 1650                                |
| muABC1 | DAIKQMKLL   | KLTKDTSADR | FLSSLGRFMA GLDTKNNVKV WFNKNGWHAI    |
| muABCR | GFLSGLGQMM  | NVSGGPVTRE | ASKEMLDFLK HLETTDN IKV WFNKNGWHAL   |
| muABCL | RTLAEIRALL  | SPQPGNALDR | ILNNLTQWAL GLDARNSLKI WFNKNGWHAM    |
|        | 1651        |            | 1700                                |
| muABC1 | SSFLNVINNA  | ILRANLQKGE | NPSQYGITAF NHPLNLTQKQ LSEVALMTTS    |
| muABCR | VSFLNVAHNA  | ILRASLPRDR | DPEEYGITVI SQPLNLTKEQ LSDITVLTTTS   |
| muABCL | VAFVNRRANNG | LLHALLPSGP | VRHAHSITTL NHPLNLTKEQ LSEATLIASS    |
|        | 1701        |            | 1750                                |
| muABC1 | VDVLVSICVI  | FAMSFVPASF | VVFLIQERVS KAKHLQFISG VKPVIYWLSN    |
| muABCR | VDAVVAICVI  | FAMSFVPASF | VLYLIQERVT KAKHLQFISG VSPTTYWLTN    |
| muABCL | VDVLVSICVV  | FAMSFVPASF | TLVLIEERIT RAKHLQLVSG LPQTLYWLG N   |
|        | 1751        |            | 1800                                |
| muABC1 | FVWDMCNYVV  | PATLVIIIFI | CFQQKSYVSS TNLPVLALLL LLYGWSITPL    |
| muABCR | FLWDIMNYAV  | SAGLVVGIFI | GFQKKAYTSP DNLPALVSL LLYGWAVIPM     |
| muABCL | FLWDMCNYLV  | AVCIVVFIFL | AFQQRAYVAP ENLPALLLL LLYGWSITPL     |
|        | 1801        |            | 1850                                |
| muABC1 | MYPASFVFKI  | PSTAYVVLTS | VNLFIGINGS VATFVLELFT NN.KLNDIND    |
| muABCR | MYPASFLFEV  | PSTAYVALSC | ANLFIGINSS AITFVLELFE NNRTLRLRFA    |
| muABCL | MYPASFFFSV  | PSTAYVVLTC | INLFIGINSS MATFVLELLS DQ.NLQEVSR    |
|        | 1851        |            | 1900                                |
| muABC1 | ILKSVFLIFP  | HFCLGRGLID | MVKNQAMADA LERFGENRFV SPLSWDLVGR    |
| muABCR | MLRKLLIVFP  | HFCLGRGLID | LALSQAVTDV YAQFGEEYSA NPFQWDLIGK    |
| muABCL | ILKQVFLIFP  | HFCLGRGLID | MVRNQAMADA FERLGDKQFQ SPLRWDIIGK    |
|        | 1901        |            | 1950                                |
| muABC1 | NLFAMAVEGV  | VFFLITVLIQ | YRFFIRPRPV KAKLPPLNDE DEDVRRERQR    |
| muABCR | NLVAMAIEGV  | VYFLLTLLIQ | HHFFLTRWIA EPAREPVFDE DDDVAEERQR    |
| muABCL | NLLAMMAQGP  | LFLITLTLQ  | HRNRLLPQSK PRLLPPLGEE DEDVAQERER    |
|        | 1951        |            | 2000                                |
| muABC1 | ILDGGGQNDI  | LEIKELTKIY | RRKRKPAVDR ICIGIPPGE C FGLLG VNGAG  |
| muABCR | VMSGGNKTDI  | LKLNELTKVY | SGSSSPAVDR LCVGV RPGE C FGLLG VNGAG |
| muABCL | VTKGATQGDV  | LVLRLDTKVY | RGQRNPAVDR LCLGIPPGE C FGLLG VNGAG  |

## FIG. 4E

|        |            |            |                                   |
|--------|------------|------------|-----------------------------------|
|        | 2001       |            | 2050                              |
| muABC1 | KSTTFKMLTG | DTPVTRGDAF | LNKNSILSNI HEVHQNMGYC PQFDAITELL  |
| muABCR | KTTTFKMLTG | DTTVTSGDAT | VAGKSILTSI SDVHQNMGYC PQFDAIDDLL  |
| muABCL | KTSTFRMVTG | DTLPSSGEAV | LAGHNVAQER SAAHRSMGYC PQSDAIFDLL  |
|        | 2051       |            | 2100                              |
| muABC1 | TGREHVEFFA | LLRGVPEKEV | GKFGEWAIK LGLVKYGEKY ASNYSGGNKR   |
| muABCR | TGREHLYLYA | RLRGVPSKEI | EKVANWGIQS LGLSLYADRL AGTYSGGNKR  |
| muABCL | TGREHLELFA | RLRGVPEAQV | AQTALSGLVR LGLPSYADRP AGTYSGGNKR  |
|        | 2101       |            | 2150                              |
| muABC1 | KLSTAMALIG | GPPVVFLDEP | TTGMDPKARR FLWNCALSIV KEGRSVVLTS  |
| muABCR | KLSTAIALTG | CPPLLLLDEP | TTGMDPQARR MLWNTIVSII REGRAVVLTS  |
| muABCL | KLATALALVG | DPAVVFLDEP | TTGMDPSARR FLWNSLLSVV REGRSVVLTS  |
|        | 2151       |            | 2200                              |
| muABC1 | HSMEECEALC | TRMAIMVNGR | FRCLGSVQHL KNRFEGDGYTI VVRIAGS... |
| muABCR | HSMEECEALC | TRLAIMVKGT | FQCLGTIQHL KYKFGDGYIV TMKIKSPKDD  |
| muABCL | HSMEECEALC | TRLAIMVNGR | FRCLGSSQHL KGRFGAGHTL TLRVPPD...  |
|        | 2201       |            | 2250                              |
| muABC1 | .NPDLPVQE  | FFGLAFPGSV | LKEKHRNMLQ YQLPSS.LSS LARIFSILSQ  |
| muABCR | LLPDLNPVEQ | FFQGNFPGSV | QRRHHSMLQ FQVPS...SS LARIFQLLIS   |
| muABCL | .Q..PEPAIA | FIRITFPGAE | LREVHGSRLR FQLPPGGRCT LTRVFRELAA  |
|        | 2251       |            | 2300                              |
| muABC1 | SKKRLHIEDY | SVSQTTLDQV | FVNFAKDQSD DDHLKDLSLH KNQTVVDVAV  |
| muABCR | HKDSSLIEEY | SVTQTTLDQV | FVNFAKQOTE TYDLPLHPRA AGASWQAKLE  |
| muABCL | QGRAHGVEDF | SVSQTTLEEV | FLYFSKDQGE EEESRQEAE EEEVSKPGRQ   |
|        | 2301       |            | 2337                              |
| muABC1 | LTSFLQDEKV | KESYV..... | .....                             |
| muABCR | EKSGRLQTQE | PLPAGSEQLA | NGSNPTAAED KHTRSPQ                |
| muABCL | HPKRVSRLF  | DPSSVETMI. | .....                             |

## FIG. 5A

|        |            |            |                                   |
|--------|------------|------------|-----------------------------------|
|        | 1          |            | 50                                |
| huABC1 | MACWPQLRLL | LWKNLTFRRR | QTCQLLLEVA WPLFIFLILI SVRLSYPPYE  |
| huABCR | MGFVRQIQLL | LWKNWTLRKR | QKIRFVVELV WPLSLFLVLI WLRNANPLYS  |
| huABCL | MAFWTQLMLL | LWKNFMYRRR | QPVQLLVELL WPLFLFFILV AVRHSHPPLE  |
|        | 51         |            | 100                               |
| huABC1 | QHECHFPNKA | MPSAGTLPWV | QGIICNANNP CFPYPTPGEA PGVVGNFNKS  |
| huABCR | HHECHFPNKA | MPSAGMLPWL | QGIFCNVNNP CFQSPTPGES PGIVSNYNNS  |
| huABCL | HHECHFPNKP | LPSAGTVPWL | QGLICNVNNT CFPQLTPGEE PGRLSNFNDS  |
|        | 101        |            | 150                               |
| huABC1 | IVARLFSDAR | RLLLYSQKDT | SMKDMRKVLR TLQQIKKSS. ....SN      |
| huABCR | ILARVYRDFQ | ELLMNAPESQ | HLGRIWTELH ILSQFMDTLR THPERIAGRG  |
| huABCL | LVSRLADAR  | TVLGGASAGR | TLAGLGKLIATLRAARSTA. ....         |
|        | 151        |            | 200                               |
| huABC1 | LKLQDFLVDN | ETFSGFLYHN | LSLPKSTVVK MLRADVILHK VFLQGYQLHL  |
| huABCR | IRIRDILKDE | ETLTLFLIKN | IGLSDSVVYL LINSQVRPEQ FAHGVPDLAL  |
| huABCL | .....      | .....      | .....                             |
|        | 201        |            | 250                               |
| huABC1 | TSLCNGSKSE | EMIQLG.... | ..DQEVSELC GLPREKLAAA ERVLRSNMDI  |
| huABCR | KDIACSEALL | ERFIIFSQR  | GAKTVRYALC SLSQGTQWI EDTLYANVDF   |
| huABCL | .....      | .....      | QPQPTKQSPL EPPMLDVAEL             |
|        | 251        |            | 300                               |
| huABC1 | LKPILRTLNS | TSPFPSKELA | EATKTLLHSL GTLAQEELFSM RSWSDMRQEV |
| huABCR | FKLFRVLPTL | LDSRSQGINL | RSWGGILSDM SPRIQEFIHR PSMQDLLWVT  |
| huABCL | LTSLLRTESL | GLALGQA..Q | EPLHSLLEAA EDLAQELLAL RSLVELRALL  |
|        | 301        |            | 350                               |
| huABC1 | MFLTNVNSSS | SSTQIYQAVS | RIVCGHPEGGLKIKSLNWEEDNNYKALFG     |
| huABCR | RPLMQNGGPE | TFTKLMGILS | DLGCGYPEGG GSRVLSFNWEEDNNYKAFLG   |
| huABCL | QRPRGTSGP. | .....LELLS | EALCSVRGPGS STVGPSLNWEASDLMELVG   |
|        | 351        |            | 400                               |
| huABC1 | GNGTEEDAET | FYDNSTTPYC | NOLMKNLESS PLSRIIWKAL KPLLVGKILY  |
| huABCR | IDSTRKDPIY | SYDRRTTSFC | NALIQSLESN PLTKIAWRRA KPLLMGKILY  |
| huABCL | ....QEPESA | LPDSSLSPAC | SELIGALDSH PLSRLWRRLL KPLILGKLLF  |
|        | 401        |            | 450                               |
| huABC1 | TPDTPATRQV | MAEVNKTQFQ | LAVFHDLEGM WEELSPKIWT FMENSQEMDL  |
| huABCR | TPDSPAARRI | LKNANSTFEE | LEHVRKLVKA WEEVGPQIWF FDNSTQNMN   |
| huABCL | APDTPFTRKL | MAQVNRTFEE | LTLLRDVREV WEMLGPRIFT FMNDSSNVAM  |
|        | 451        |            | 500                               |
| huABC1 | VRMLLDSDRN | DHFEQQQLDG | LDWTAQDIVA FLAKHPEDVQ SSNGSVYTWR  |
| huABCR | IRDTLGNPTV | KDFLNRQLGE | EGITAEAILN FLYKGPRESQ ADDMANFDWR  |
| huABCL | LQRLQLMQDE | GRRQPRPGGR | D.....HMEALRSFL DPGSGGYSWQ        |

## FIG. 5B

|        |            |            |                                   |
|--------|------------|------------|-----------------------------------|
|        | 501        |            | 550                               |
| huABC1 | EAFNETNQAI | RTISRFMECV | NLNKLEPIAT EVWLINKSME LLDERKFWAG  |
| huABCR | DIFNITDRTL | RLVNQYLECL | VLDKFESYND ETQLTQRALS LLEENMFWAG  |
| huABCL | DAHADVGLV  | GTLGRVTECL | SLDKLEAAPS EAALVSRALQ LLAEHRFWAG  |
|        | 551        |            | 600                               |
| huABC1 | IVFTG..... | ...ITPGSIE | LPHHVKYKIR MDIDNVERTN KIKDGYWDPG  |
| huABCR | VVFPD..... | ...MYPWTSS | LPPHVKYKIR MDIDVVEKTN KIKDRYWDSG  |
| huABCL | VVFLGPEDSS | DPTEHPTPDL | GPGHVRIKIR MDIDVVTRTN KIRDREWDPG  |
|        | 601        |            | 650                               |
| huABC1 | PRADPFEDMR | YVWGGFAYLQ | DVVEQAIIRV LTGTEKKTGV YMQQMPYPCY  |
| huABCR | PRADPVEDFR | YIWGGFAYLQ | DMVEQGITRS QVQAEAPVGI YLQQMPYPCF  |
| huABCL | PAADPLTDLR | YVWGGFVYLQ | DLVERAAVRV LSGANPRAGL YLQQMPYPCY  |
|        | 651        |            | 700                               |
| huABC1 | VDDIFLRVMS | RSMPLFMTLA | WIYSVAVIK GIVYEKEARL KETMRIMGLD   |
| huABCR | VDDSFMIILN | RCFPIFMVLA | WIYSVSMTVK SIVLEKELRL KETLKNQGV   |
| huABCL | VDDVFLRVLS | RSLPLFLTLA | WIYSVTLTVK AVVREKETRL RDTMRAMGLS  |
|        | 701        |            | 750                               |
| huABC1 | NSILWFSWFI | SSLIPLLVA  | GLLVVILKLG NLLPYSDPSV VVFVLSVFAV  |
| huABCR | NAVIWCTWFL | DSFSIMSMI  | FLLTIFIMHG RILHYSDPFI LFLFLAFST   |
| huABCL | RAVLWLGWFL | SCLGPFLLSA | ALLVLVLKLG DILPYSHPGV VFLFLAAFAV  |
|        | 751        |            | 800                               |
| huABC1 | VTILQCFLIS | TLFSRANLAA | ACGGIIYFTL YLPYVLCVAW QDYVGFTLKI  |
| huABCR | ATIMLCFLLS | TFFSKASLAA | ACSGVIYFTL YLPHILCFAW QDRMTAELKK  |
| huABCL | ATVTQSFLLS | AFFSRANLAA | ACGGLAYFSL YLPYVLCVAW RDRLPAGGRV  |
|        | 801        |            | 850                               |
| huABC1 | FASLLSPVAF | GFGCEYFALF | EEQGIGVQWD NLFESPVEED GFNLTSVSM   |
| huABCR | AVSLLSPVAF | GFGTEYLVRF | EEQGLGLQWS NIGNSPTEGD EFSFLLSMQM  |
| huABCL | AASLLSPVAF | GFGCESLALL | EEQGEQAQWH NVG.TRPTAD VFSLAQVSGL  |
|        | 851        |            | 900                               |
| huABC1 | MLFDTFLYGV | MTWYIEAVFP | GQYGIPRPWY FPCTKSYWFG .....       |
| huABCR | MLLDAAVYGL | LAWYLDQVFP | GDYGTPLPWY FLLQESYWLS GEGCSTREER  |
| huABCL | LLDAAALYGL | ATWYLEAVCP | GQYGIPEPWN FPFRRSYWCG .....       |
|        | 901        |            | 950                               |
| huABC1 | .....EESDE | KSHPGSNQKR | ISEICMEEEP THLKLGVSIQ NLVKVYRDGM  |
| huABCR | ALEKTEPLTE | ETEDPEHPEG | IHDSFFEREH PGWVPGVCVK NLVKIFEPCG  |
| huABCL | .....PRPP  | KSPAPCPTPL | DPKVLVEEAP PGLSPGVSVR SLEKRFPGSP  |
|        | 951        |            | 1000                              |
| huABC1 | KVAVDGLALN | FYEGQITSFL | GHNGAGKTTT MSILTGLFPP TSGTAYILGK  |
| huABCR | RPAVDRLNIT | FYENQITAF  | LGHNGAGKTTT LSILTGLLPP TSGTVLVGGR |
| huABCL | QPALRGLSLD | FYQGHITAF  | LGHNGAGKTTT LSILSGLFPP SGGSAFILGH |

## FIG. 5C

|        |             |             |                                    |
|--------|-------------|-------------|------------------------------------|
|        | 1001        |             | 1050                               |
| huABC1 | DIRSEMSTIR  | QNLGVCPQHN  | VLFDMLTVEE HIWFIYARLKG LSEKHVKAEM  |
| huABCR | DIETSLDAVR  | QSLGMCPQHN  | ILFHHLTVAE HMLFYAQLKG KSQEEAQLEM   |
| huABCL | DVRSSMAAIR  | PHLGVCPQYN  | VLFDMLTVDE HVWFIYGRLLKG LSAAVVGPEQ |
|        | 1051        |             | 1100                               |
| huABC1 | EQMALDVGLP  | SSKLKSKTSQ  | LSGGMQRKLS VALAFVGGSK VVILDEPTAG   |
| huABCR | EAMLEDTGLH  | HK.RNEEAQD  | LSGGMQRKLS VAIAFVGDAK VVILDEPTSG   |
| huABCL | DRLLQDVGLV  | SK.QSVQTRH  | LSGGMQRKLS VAIAFVGGSQ VVILDEPTAG   |
|        | 1101        |             | 1150                               |
| huABC1 | VDPYSRRGIW  | ELLLKYRQGR  | TIILSTHHMD EADVLGDRIA IISHGKLCCV   |
| huABCR | VDPYSRRSIW  | DLLLKYRSGR  | TIIMSTHHMD EADLLGDRIA IIAQGRLYCS   |
| huABCL | VDPASRRGIW  | ELLLKYREGR  | TLILSTHHL D EAELLGDRVA VVAGGRLCCC  |
|        | 1151        |             | 1200                               |
| huABC1 | GSSLFLKNQL  | GTGYLTLVK   | KDVESSLSSC RNSSSTVSYL KKEDSVSQSS   |
| huABCR | GTPLFLKNCF  | GTGLYTLVR   | KMKNIQSQRK ....GSEGT C SCSSKGFSTT  |
| huABCL | GSPLFLRRHL  | GSGYYLTLVK  | ARLPLTTNEK ..... ..ADTMEGS         |
|        | 1201        |             | 1250                               |
| huABC1 | SDAGLGSDHE  | SDTLTIDVSA  | ISNLIRKHVS EARLVEDIGH ELTYVLPYEA   |
| huABCR | CPAHVDDLTP  | EQVLDGDVNE  | LMDVVLHHVP EAKLVECIGQ ELIFLLPNKN   |
| huABCL | VDTRQEKKNQ  | SQGSRVGTPTQ | LLALVQHWVP GARLVEELPH ELVLVLPYTG   |
|        | 1251        |             | 1300                               |
| huABC1 | AKEGAFVELF  | HEIDRLSDL   | GISSYGISET TLEEIFLKVA EESGVDAETS   |
| huABCR | FKHRAYASLF  | RELEETLADL  | GLSSFGISDT PLEEIFLKVT EDSDSGPLFA   |
| huABCL | AHDGSFATLF  | RELDTRLAEL  | RLTGYGISDT SLEEIFLKVV EECAADTME    |
|        | 1301        |             | 1350                               |
| huABC1 | DGTLPARRRNR | RAFGDKQSCL  | RPFTEDDAAD PNDSIDIPES RETDLLSGMD   |
| huABCR | GGAQQKREN.  | ..VNPRHPCL  | GPREKAGQTP QDSNVCPGA PAAHPEGQPP    |
| huABCL | DGSCGQHLCT  | .GIAGLDVTL  | RLKMPPQETA LENGEPAGSA PETDQSGSPD   |
|        | 1351        |             | 1400                               |
| huABC1 | GK....GSYQ  | VKGWKLTOQQ  | FVALLWKRLI IARRSRKGF A QIVLPAVFV   |
| huABCR | PEPECPPGQL  | NTGTQLVLQH  | VQALLVKRFQ HTIRSHKDFL A QIVLPATFV  |
| huABCL | AV....G..R  | VQGWALTRQQ  | LQALLKRFL LARRSRGLF A QIVLPALFV    |
|        | 1401        |             | 1450                               |
| huABC1 | CIALVFSLIV  | PPFGKYPSE   | LQPMWYNEQY TFVSNDAPE TGTLELLNAL    |
| huABCR | FLALMLSIVI  | LPFGEYPALT  | LHPWIYQQY TFFSMDEPGS EQFTVLADVL    |
| huABCL | GLALVFSLIV  | PPFGHYPALR  | LSPTMYGAQV SFFSEDAPGD PGRARLLEAL   |
|        | 1451        |             | 1500                               |
| huABC1 | TKDPGFGTRC  | MEGNPIPDTP  | CQAGEEEWTT APVPQTIMDL FQNGNWTMQN   |
| huABCR | LKNPGFGNRC  | LKEGWLPEYP  | CGN.STPWKT PSVSPNITQL FQKQKWTQVN   |
| huABCL | L.....      | .QEAGLEPP   | VQHSSHRFSA PEVPAEVAKV LASGNWTPES   |

## FIG. 5D

|        |            |             |                                    |
|--------|------------|-------------|------------------------------------|
|        | 1501       |             | 1550                               |
| huABC1 | PSPACQCSSD | KIKKMLPVCP  | PGAGGLPPPQ RKQNTADILQ DLTGRNISDY   |
| huABCR | PSPSCRCSTR | EKLTMLEPCP  | EGAGGLPPPQ RTQRSTEILQ DLTDRNISDF   |
| huABCL | PSPACQCSQP | GARRLLPDCP  | AAAGGPPPPQ AVTGSGEVVQ NLTGRNLSDF   |
|        | 1551       |             | 1600                               |
| huABC1 | LVKTYVQIIA | KSLKNKIWVN  | EFYGGFSLG VSNTQALPPS QEVNDATKQM    |
| huABCR | LVKTYPALIR | SSLKSKFWVN  | EQRYGGISIG GKLPVVPITG EALVGFLSDL   |
| huABCL | LVKTYPRLVR | QGLKTKKWVN  | EVRYGGFSLG GRDPGLP.SG QELGRSVEEL   |
|        | 1601       |             | 1650                               |
| huABC1 | KKHLKLAKDS | SADRFLNSLG  | RFMTGLDTRN NVKVFNNKG WHAISSFLNV    |
| huABCR | GRIMNVSGGP | ITREASKEIP  | DFLKHLETED NIKVWFNNKG WHALVSFLNV   |
| huABCL | WALLSPLPGG | ALDRVLKNLT  | AWAHS LDAQD SLKIWFNNKG WHSMVAFVNR  |
|        | 1651       |             | 1700                               |
| huABC1 | INNAILRANL | QKGENPSHYG  | ITAFNHPLNL TKQQLSEVAP MTTSDVDLVS   |
| huABCR | AHNAILRASL | PKDRSPEEYG  | ITVISQPLNL TKEQLSEITV LTTSVDAVVA   |
| huABCL | ASNAILRAHL | PPGPARRAHS  | ITTLNHPLNL TKEQLSEAL MASSVDLVS     |
|        | 1701       |             | 1750                               |
| huABC1 | ICVIFAMSFV | PASFVVFLIQ  | ERVSKAKHLQ FISGVKPIY WLSNFVWDMC    |
| huABCR | ICVIFSMSFV | PASFVLYLIQ  | ERVNKS KHLQ FISGVSPPTY WVTNFWLWDIM |
| huABCL | ICVVFAMSFV | PASFTLV LIE | ERVTRAKHLQ LMGGLSPTLY WLG NFWLWDMC |
|        | 1751       |             | 1800                               |
| huABC1 | NYVVPATLVI | IIFICFQOKS  | YVSSTNLPVL ALLLLLYGWS ITPLMPASF    |
| huABCR | NYSVSAGLVV | GIFIGFQKKA  | YTSPENLPAL VALLLLYGWA VIPMMYPASF   |
| huABCL | NYLVPACIVV | LIFLAFQORA  | YVAPANLPAL LLLLLLYGWS ITPLMPASF    |
|        | 1801       |             | 1850                               |
| huABC1 | VFKIPSTAYV | VLTSVNLFIG  | INGSVATFVL ELFTDN.KLN NINDILKSVE   |
| huABCR | LFDVPSTAYV | ALSCANLFIG  | INSSAITFIL ELFDNNRTL RFNAVLRKLL    |
| huABCL | FFSVPSTAYV | VLTCINLFIG  | INGSMATFVL ELFS DQ.KLQ EVSRILKQVF  |
|        | 1851       |             | 1900                               |
| huABC1 | LIFPHFCLGR | GLIDMVKNQA  | MADALERFGE NRFVSPLSWD LVGRNLFAMA   |
| huABCR | IVFPHFCLGR | GLIDLALSQA  | VTDVYARFGE EHSANPFHWD LIGKNLFAMV   |
| huABCL | LIFPHFCLGR | GLIDMVRNQA  | MADAFERLGD RQFQSPLRWE VVGKNLLAMV   |
|        | 1901       |             | 1950                               |
| huABC1 | VEGVVFFLIT | VLIQYRFFIR  | PRPVNAKLSP LNDEDEDVRR ERQRILDGGG   |
| huABCR | VEGVVYFLLT | LLVQRHFFLS  | QWIAEPTKEP IVDEDDDAE ERQRIITGGN    |
| huABCL | IQGPLFLLFT | LLLQHRSQLL  | PQPRVRSPLP LGEEDEDVAR ERERVVQGAT   |
|        | 1951       |             | 2000                               |
| huABC1 | QNDILEIKEL | TKIYRRKRKP  | AVDRICVGIP PGECFGLLG V NGAGKSSTFK  |
| huABCR | KTDILRLHEL | TKIYLGTS SP | AVDRLCVGVR PGECFGLLG V NGAGKTTTFK  |
| huABCL | QGDVLVLRNL | TKVYRGQ RMP | AVDRLCLGIP PGECFGLLG V NGAGKTSTFR  |

## FIG. 5E

|        |             |             |             |
|--------|-------------|-------------|-------------|
|        | 2001        |             | 2050        |
| huABC1 | MLTGDTTVTR  | GDAFLNRNSI  | LSNIHEVHQN  |
| huABCR | MLTGDTTVTS  | GDATEVAGKSI | LTNISEVHQN  |
| huABCL | MVTGDTLASR  | GEAVLAGHSV  | AREPSAAHLS  |
|        |             |             | MGYCPQFDAI  |
|        |             |             | TELLTGREHV  |
|        |             |             | DELLTGREHL  |
|        |             |             | FELLTGREHL  |
|        | 2051        |             | 2100        |
| huABC1 | EFFALLRGVP  | EKEVGKVGW   | AIRKLGLVKY  |
| huABCR | YLYARLRGVP  | AEEIEKVANW  | SIKSLGLTVY  |
| huABCL | ELLARLRGVP  | EAQVAQTAS   | GLARLGLSWY  |
|        |             |             | ADCLAGTYS   |
|        |             |             | GNKRKLSTAM  |
|        |             |             | GNKRKLSTAI  |
|        |             |             | GNKRKLATAL  |
|        | 2101        |             | 2150        |
| huABC1 | ALIGGPPVVF  | LDEPTTGMDP  | KARRFLWNCA  |
| huABCR | ALIGCPPLVL  | LDEPTTGMDP  | QARRMLWNVI  |
| huABCL | ALVGDPVAVF  | LDEPTTGMDP  | SARRFLWNSL  |
|        |             |             | LSVVKEGRSV  |
|        |             |             | VLTSHSMEEC  |
|        |             |             | VLTSHSMEEC  |
|        |             |             | VLTSHSMEEC  |
|        | 2151        |             | 2200        |
| huABC1 | EALCTRMAIM  | VNGRFRCLGS  | VQHLKNRFGD  |
| huABCR | EALCTRMAIM  | VKGAFRCMGT  | IQHLKSKFGD  |
| huABCL | EALCSRMAIM  | VNGRFRCLGS  | PQHLKGRFAA  |
|        |             |             | GYTIVVRIAG  |
|        |             |             | S....NPDLK  |
|        |             |             | PKDDLPLDLN  |
|        |             |             | A.....RSQ   |
|        | 2201        |             | 2250        |
| huABC1 | PVQDFFGGLAF | PGSVPKEKHR  | NMLQYQLPSS  |
| huABCR | PVEQFFQGNF  | PGSVQRRHY   | NMLQFQVSS   |
| huABCL | PAAAFVAAEF  | PGSELREAHG  | GRLRFQLPPG  |
|        |             |             | .LSSLARIFS  |
|        |             |             | ILSQSKKRLH  |
|        |             |             | LLLSHKDSL   |
|        |             |             | ELAVHGAEHG  |
|        | 2251        |             | 2300        |
| huABC1 | IEDYSVSQTT  | LDQVFVNFAK  | DQSDDDHLKD  |
| huABCR | IEEYSVTQTT  | LDQVFVNFAK  | QQTESHDLPL  |
| huABCL | VEDFSVSQTM  | LEEVFLYFSK  | DQGKDEEDTEE |
|        |             |             | LSLHKNQTVV  |
|        |             |             | DVAVLTSFLQ  |
|        |             |             | AQD.....    |
|        |             |             | APGLQHPKRV  |
|        | 2301        |             |             |
| huABC1 | DEKVKEYSV   |             | ....        |
| huABCR | .....       |             | ....        |
| huABCL | SQFLDDPSTA  | ETVL        |             |

## SEQUENCE LISTING

<110> Shutter, John  
Ulias, Laarni

<120> ATP-Binding Cassette Transporter-Like Molecules and  
Uses Thereof

<130> 00-658-B

<140>

<141>

<150> 60/253,520

<151> 2000-11-28

<160> 24

<170> PatentIn Ver. 2.0

<210> 1

<211> 6633

<212> DNA

<213> Mus musculus

<220>

<221> sig\_peptide

<222> (1)..(138)

<220>

<221> CDS

<222> (1)..(6504)

<400> 1

|                                                                 |    |
|-----------------------------------------------------------------|----|
| atg gct ttc tgc aca cag ttg atg ctt ctg ctg tgg aaa aat tac acc | 48 |
| Met Ala Phe Cys Thr Gln Leu Met Leu Leu Leu Trp Lys Asn Tyr Thr |    |
| 1 5 10 15                                                       |    |

|                                                                 |    |
|-----------------------------------------------------------------|----|
| tat cga cgg aga caa ccg atc caa cta cta gtg gag ttg ctt tgg ccc | 96 |
| Tyr Arg Arg Arg Gln Pro Ile Gln Leu Leu Val Glu Leu Leu Trp Pro |    |
| 20 25 30                                                        |    |

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| ctc ttc ctc ttc ttc atc cta gtg gct gtc cgt cac tcc cac ccc cct | 144 |
| Leu Phe Leu Phe Phe Ile Leu Val Ala Val Arg His Ser His Pro Pro |     |
| 35 40 45                                                        |     |

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| ctg gag cat cac gaa tgc cac ttt cca aac aag cca tta cca tcg gcg | 192 |
| Leu Glu His His Glu Cys His Phe Pro Asn Lys Pro Leu Pro Ser Ala |     |
| 50 55 60                                                        |     |

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| ggc acg gtg ccc tgg ctg cag ggc ctt gtc tgc aac gta aac aac tcc | 240 |
| Gly Thr Val Pro Trp Leu Gln Gly Leu Val Cys Asn Val Asn Asn Ser |     |
| 65 70 75 80                                                     |     |

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| tgc ttc cag cac cca acg cct ggc gag aag cct ggg gtc ctg agt aac | 288 |
| Cys Phe Gln His Pro Thr Pro Gly Glu Lys Pro Gly Val Leu Ser Asn |     |
| 85 90 95                                                        |     |

|                                                                 |      |
|-----------------------------------------------------------------|------|
| ttt aag gat tcc ttg atc tcg agg ctc ctc gct gat acc cgc aca gtg | 336  |
| Phe Lys Asp Ser Leu Ile Ser Arg Leu Leu Ala Asp Thr Arg Thr Val |      |
| 100 105 110                                                     |      |
| ctc ggg ggc cac agc atc cag gac atg ctg gat gcc ctg ggg aaa ctg | 384  |
| Leu Gly Gly His Ser Ile Gln Asp Met Leu Asp Ala Leu Gly Lys Leu |      |
| 115 120 125                                                     |      |
| atc ccc gtg ctc agg gca gtt gga ggt gga gca cga cca cag gag agt | 432  |
| Ile Pro Val Leu Arg Ala Val Gly Gly Gly Ala Arg Pro Gln Glu Ser |      |
| 130 135 140                                                     |      |
| gac cag ccg acc agt caa ggg tca gtg act aag ctt ctg gag aag atc | 480  |
| Asp Gln Pro Thr Ser Gln Gly Ser Val Thr Lys Leu Leu Glu Lys Ile |      |
| 145 150 155 160                                                 |      |
| ctg caa agg gca tcc ctg gat cct gtg ctg ggt caa gcc cag gat tct | 528  |
| Leu Gln Arg Ala Ser Leu Asp Pro Val Leu Gly Gln Ala Gln Asp Ser |      |
| 165 170 175                                                     |      |
| atg aga aag ttc tca gat gct atc agg gat ctt gcc cag gag ctc ctg | 576  |
| Met Arg Lys Phe Ser Asp Ala Ile Arg Asp Leu Ala Gln Glu Leu Leu |      |
| 180 185 190                                                     |      |
| aca ctg ccc agc ctg atg gag ctc cga gct ttg ctg cgg agg ccc cga | 624  |
| Thr Leu Pro Ser Leu Met Glu Leu Arg Ala Leu Leu Arg Arg Pro Arg |      |
| 195 200 205                                                     |      |
| ggg tca gct ggt tct ctg gag ctg gtt tcg gag gcc ctc tgc agt acc | 672  |
| Gly Ser Ala Gly Ser Leu Glu Leu Val Ser Glu Ala Leu Cys Ser Thr |      |
| 210 215 220                                                     |      |
| aag gga ccc agc agt cca ggg ggc ctg tcc ctc aat tgg tac gaa gcc | 720  |
| Lys Gly Pro Ser Ser Pro Gly Gly Leu Ser Leu Asn Trp Tyr Glu Ala |      |
| 225 230 235 240                                                 |      |
| aac cag ctt aat gag ttc atg ggg cca gag gtg gcc cct gcc ctg cct | 768  |
| Asn Gln Leu Asn Glu Phe Met Gly Pro Glu Val Ala Pro Ala Leu Pro |      |
| 245 250 255                                                     |      |
| gac aac agt ctc agc cct gcc tgc tct gag ttt gtg ggg aca ctg gat | 816  |
| Asp Asn Ser Leu Ser Pro Ala Cys Ser Glu Phe Val Gly Thr Leu Asp |      |
| 260 265 270                                                     |      |
| gac cac cct gtg tct cgg ctg ctc tgg agg cgc ctg aag cca ttg atc | 864  |
| Asp His Pro Val Ser Arg Leu Leu Trp Arg Arg Leu Lys Pro Leu Ile |      |
| 275 280 285                                                     |      |
| ctc ggg aaa att ctc ttt gca cct gac aca aac ttc act cgg aag ctc | 912  |
| Leu Gly Lys Ile Leu Phe Ala Pro Asp Thr Asn Phe Thr Arg Lys Leu |      |
| 290 295 300                                                     |      |
| atg gct cag gtg aac cag acc ttc gag gag ctg gct ctg ttg agg gac | 960  |
| Met Ala Gln Val Asn Gln Thr Phe Glu Glu Leu Ala Leu Leu Arg Asp |      |
| 305 310 315 320                                                 |      |
| cta cac gaa ctc tgg ggg gtg ctg gga ccc cag atc ttc aac ttc atg | 1008 |
| Leu His Glu Leu Trp Gly Val Leu Gly Pro Gln Ile Phe Asn Phe Met |      |

3/93

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| ctc | act | gtg | aag | gcc | gtg | gtg | cgt | gag | aaa | gag | aca | cgg | ctg | cga | gaa | 1728 |
| Leu | Thr | Val | Lys | Ala | Val | Val | Arg | Glu | Lys | Glu | Thr | Arg | Leu | Arg | Glu |      |
|     |     |     | 565 |     |     |     |     |     | 570 |     |     |     |     |     | 575 |      |
| acc | atg | cgt | gcg | atg | ggg | ctg | agc | cgc | gcg | gtg | ctc | tgg | ctt | ggt | tgg | 1776 |
| Thr | Met | Arg | Ala | Met | Gly | Leu | Ser | Arg | Ala | Val | Leu | Trp | Leu | Gly | Trp |      |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |      |
| ttc | ctc | agc | tgc | ctg | gga | ccc | ttc | ctg | gtc | agc | gct | gcg | ttg | ctg | gta | 1824 |
| Phe | Leu | Ser | Cys | Leu | Gly | Pro | Phe | Leu | Val | Ser | Ala | Ala | Leu | Leu | Val |      |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |      |
| tta | gtg | ctt | aag | cta | ggg | aac | atc | ctt | cct | tac | agc | cac | ccg | ggt | gta | 1872 |
| Leu | Val | Leu | Lys | Leu | Gly | Asn | Ile | Leu | Pro | Tyr | Ser | His | Pro | Val | Val |      |
|     | 610 |     |     |     |     | 615 |     |     |     |     | 620 |     |     |     |     |      |
| atc | ttc | ctt | ttc | ttg | gcg | gcc | ttc | gcg | gtg | gcc | acc | gtc | gct | cag | agt | 1920 |
| Ile | Phe | Leu | Phe | Leu | Ala | Ala | Phe | Ala | Val | Ala | Thr | Val | Ala | Gln | Ser |      |
| 625 |     |     |     |     | 630 |     |     |     |     | 635 |     |     |     |     | 640 |      |
| ttt | ctg | ctc | agc | gcc | ttc | ttc | tcc | agg | gcc | aat | ctg | gca | gca | gcc | tgc | 1968 |
| Phe | Leu | Leu | Ser | Ala | Phe | Phe | Ser | Arg | Ala | Asn | Leu | Ala | Ala | Ala | Cys |      |
|     |     |     | 645 |     |     |     |     |     | 650 |     |     |     |     | 655 |     |      |
| ggg | ggg | ctc | gcc | tat | ttt | gcg | ctc | tat | ctg | ccc | tac | gta | ctg | tgt | gtg | 2016 |
| Gly | Gly | Leu | Ala | Tyr | Phe | Ala | Leu | Tyr | Leu | Pro | Tyr | Val | Leu | Cys | Val |      |
|     |     |     | 660 |     |     |     |     | 665 |     |     |     |     | 670 |     |     |      |
| gcc | tgg | cgc | gag | cgc | ctg | cac | ctg | ggc | gga | ctc | tta | gct | gcg | agc | ctg | 2064 |
| Ala | Trp | Arg | Glu | Arg | Leu | His | Leu | Gly | Gly | Leu | Leu | Ala | Ala | Ser | Leu |      |
|     |     | 675 |     |     |     |     | 680 |     |     |     |     | 685 |     |     |     |      |
| ctg | tcc | cct | gta | gcc | ttt | ggc | ttt | gga | tgc | gaa | agc | ctg | gcg | cta | cta | 2112 |
| Leu | Ser | Pro | Val | Ala | Phe | Gly | Phe | Gly | Cys | Glu | Ser | Leu | Ala | Leu | Leu |      |
|     | 690 |     |     |     |     | 695 |     |     |     |     | 700 |     |     |     |     |      |
| gag | gag | cag | gga | gac | ggg | gct | cag | tgg | cac | aat | ttg | ggc | aca | ggc | ccc | 2160 |
| Glu | Glu | Gln | Gly | Asp | Gly | Ala | Gln | Trp | His | Asn | Leu | Gly | Thr | Gly | Pro |      |
| 705 |     |     |     |     | 710 |     |     |     |     | 715 |     |     |     |     | 720 |      |
| gcg | gag | gac | gtc | ttc | agc | ctg | gcc | cag | gtg | tct | gcc | ttc | ctg | ttg | ctt | 2208 |
| Ala | Glu | Asp | Val | Phe | Ser | Leu | Ala | Gln | Val | Ser | Ala | Phe | Leu | Leu | Leu |      |
|     |     |     | 725 |     |     |     |     |     | 730 |     |     |     |     | 735 |     |      |
| gat | gcc | gtc | atc | tac | ggc | ctt | gcc | ctc | tgg | tac | cta | gag | gct | gtg | tgc | 2256 |
| Asp | Ala | Val | Ile | Tyr | Gly | Leu | Ala | Leu | Trp | Tyr | Leu | Glu | Ala | Val | Cys |      |
|     |     |     | 740 |     |     |     |     | 745 |     |     |     |     | 750 |     |     |      |
| cca | ggc | cag | tat | gga | atc | cct | gaa | cca | tgg | aat | ttc | cct | ttt | cgg | agg | 2304 |
| Pro | Gly | Gln | Tyr | Gly | Ile | Pro | Glu | Pro | Trp | Asn | Phe | Pro | Phe | Arg | Arg |      |
|     |     | 755 |     |     |     |     | 760 |     |     |     |     | 765 |     |     |     |      |
| agc | tac | tgg | tgt | gga | cct | ggg | cct | ccc | aag | agt | tct | gtc | ttg | gcc | cct | 2352 |
| Ser | Tyr | Trp | Cys | Gly | Pro | Gly | Pro | Pro | Lys | Ser | Ser | Val | Leu | Ala | Pro |      |
|     | 770 |     |     |     |     | 775 |     |     |     |     | 780 |     |     |     |     |      |
| gcc | cca | caa | gat | ccc | aag | gtt | ctg | gtg | gaa | gag | cca | ccc | ctt | ggc | ctg | 2400 |
| Ala | Pro | Gln | Asp | Pro | Lys | Val | Leu | Val | Glu | Glu | Pro | Pro | Leu | Gly | Leu |      |

|     |     |     |     |      |     |     |     |     |      |     |     |     |     |      |     |      |
|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|------|
| 785 |     |     |     | 790  |     |     |     | 795 |      |     |     | 800 |     |      |     |      |
| gtt | cct | ggt | gtc | tcc  | att | cga | ggc | ctg | aag  | aaa | cat | ttt | cgt | ggc  | tgt | 2448 |
| Val | Pro | Gly | Val | Ser  | Ile | Arg | Gly | Leu | Lys  | Lys | His | Phe | Arg | Gly  | Cys |      |
|     |     |     |     | 805  |     |     |     |     | 810  |     |     |     |     | 815  |     |      |
| ccg | cag | cca | gcc | ctg  | caa | gga | ctc | aac | ctt  | gac | ttc | tac | gaa | ggc  | cac | 2496 |
| Pro | Gln | Pro | Ala | Leu  | Gln | Gly | Leu | Asn | Leu  | Asp | Phe | Tyr | Glu | Gly  | His |      |
|     |     |     |     | 820  |     |     |     |     | 825  |     |     |     |     | 830  |     |      |
| atc | act | gcc | ttt | ttg  | ggt | cac | aac | ggg | gct  | ggc | aag | aca | acc | aca  | ctg | 2544 |
| Ile | Thr | Ala | Phe | Leu  | Gly | His | Asn | Gly | Ala  | Gly | Lys | Thr | Thr | Thr  | Leu |      |
|     |     |     |     | 835  |     |     |     |     | 840  |     |     |     |     | 845  |     |      |
| tcc | att | ttg | agt | ggt  | ctc | ttc | cca | ccc | agt  | agt | ggc | tcg | gcc | tcc  | atc | 2592 |
| Ser | Ile | Leu | Ser | Gly  | Leu | Phe | Pro | Pro | Ser  | Ser | Gly | Ser | Ala | Ser  | Ile |      |
|     |     |     |     | 850  |     |     |     |     | 855  |     |     |     |     | 860  |     |      |
| ctg | ggc | cat | gat | gta  | caa | acc | aac | atg | gca  | gcc | atc | cgg | ccc | cac  | ctg | 2640 |
| Leu | Gly | His | Asp | Val  | Gln | Thr | Asn | Met | Ala  | Ala | Ile | Arg | Pro | His  | Leu |      |
|     |     |     |     | 865  |     |     |     |     | 870  |     |     |     |     | 875  |     |      |
| ggc | atc | tgc | ccg | cag  | tac | aat | gtg | ctg | ttt  | gat | atg | ctg | aca | gtg  | gaa | 2688 |
| Gly | Ile | Cys | Pro | Gln  | Tyr | Asn | Val | Leu | Phe  | Asp | Met | Leu | Thr | Val  | Glu |      |
|     |     |     |     | 885  |     |     |     |     | 890  |     |     |     |     | 895  |     |      |
| gaa | cat | gtt | tgg | ttc  | tat | ggc | cgt | ttg | aaa  | ggc | gtg | agt | gca | gcc  | gcc | 2736 |
| Glu | His | Val | Trp | Phe  | Tyr | Gly | Arg | Leu | Lys  | Gly | Val | Ser | Ala | Ala  | Ala |      |
|     |     |     |     | 900  |     |     |     |     | 905  |     |     |     |     | 910  |     |      |
| atg | ggc | ccc | gag | cgg  | gaa | cgt | ctg | ata | cgg  | gat | gtg | ggg | ctt | acc  | ctc | 2784 |
| Met | Gly | Pro | Glu | Arg  | Glu | Arg | Leu | Ile | Arg  | Asp | Val | Gly | Leu | Thr  | Leu |      |
|     |     |     |     | 915  |     |     |     |     | 920  |     |     |     |     | 925  |     |      |
| aag | cgg | gac | aca | cag  | aca | cgc | cac | ctc | tct  | ggt | gga | atg | cag | aga  | aaa | 2832 |
| Lys | Arg | Asp | Thr | Gln  | Thr | Arg | His | Leu | Ser  | Gly | Gly | Met | Gln | Arg  | Lys |      |
|     |     |     |     | 930  |     |     |     |     | 935  |     |     |     |     | 940  |     |      |
| ctt | tct | gtg | gcc | att  | gcc | ttt | gtg | ggt | ggc  | tct | cgt | gtg | gtc | atc  | atg | 2880 |
| Leu | Ser | Val | Ala | Ile  | Ala | Phe | Val | Gly | Gly  | Ser | Arg | Val | Val | Ile  | Met |      |
|     |     |     |     | 945  |     |     |     |     | 950  |     |     |     |     | 955  |     |      |
| gac | gag | ccc | act | gct  | ggt | gtg | gac | ccc | gct  | tcc | cgc | cgt | ggc | att  | tgg | 2928 |
| Asp | Glu | Pro | Thr | Ala  | Gly | Val | Asp | Pro | Ala  | Ser | Arg | Arg | Gly | Ile  | Trp |      |
|     |     |     |     | 965  |     |     |     |     | 970  |     |     |     |     | 975  |     |      |
| gaa | ttg | cta | ctt | aag  | tac | aga | gaa | ggt | cgg  | aca | ctg | att | ctc | tcc  | act | 2976 |
| Glu | Leu | Leu | Leu | Lys  | Tyr | Arg | Glu | Gly | Arg  | Thr | Leu | Ile | Leu | Ser  | Thr |      |
|     |     |     |     | 980  |     |     |     |     | 985  |     |     |     |     | 990  |     |      |
| cac | cac | ctg | gat | gag  | gca | gag | ctc | ttg | gga  | gat | cgc | gtg | gcc | atg  | gtg | 3024 |
| His | His | Leu | Asp | Glu  | Ala | Glu | Leu | Leu | Gly  | Asp | Arg | Val | Ala | Met  | Val |      |
|     |     |     |     | 995  |     |     |     |     | 1000 |     |     |     |     | 1005 |     |      |
| gca | ggt | ggc | tct | ttg  | tgc | tgc | tgt | ggg | tcc  | ccg | ctt | ttc | ttg | cgc  | cga | 3072 |
| Ala | Gly | Gly | Ser | Leu  | Cys | Cys | Cys | Gly | Ser  | Pro | Leu | Phe | Leu | Arg  | Arg |      |
|     |     |     |     | 1010 |     |     |     |     | 1015 |     |     |     |     | 1020 |     |      |

6/93

| 1250                                                            | 1255 | 1260 |      |
|-----------------------------------------------------------------|------|------|------|
| ctg ttc ttc agc ctc att gtg cct cct ttt ggc cag tac cca ccc ctg |      |      | 3840 |
| Leu Phe Phe Ser Leu Ile Val Pro Pro Phe Gly Gln Tyr Pro Pro Leu |      |      |      |
| 1265                                                            | 1270 | 1275 | 1280 |
| cag ctc agc cct gct atg tat ggc cct cag gtc tcg ttc ttc agt gag |      |      | 3888 |
| Gln Leu Ser Pro Ala Met Tyr Gly Pro Gln Val Ser Phe Phe Ser Glu |      |      |      |
| 1285                                                            | 1290 |      | 1295 |
| gat gcc cct ggg gac ccc aac cgg atg aag ctg ctg gag gct ctg cta |      |      | 3936 |
| Asp Ala Pro Gly Asp Pro Asn Arg Met Lys Leu Leu Glu Ala Leu Leu |      |      |      |
| 1300                                                            | 1305 |      | 1310 |
| ggg gag gct ggg ctg cag gaa ccc agt atg cag gac aaa gat gcc agg |      |      | 3984 |
| Gly Glu Ala Gly Leu Gln Glu Pro Ser Met Gln Asp Lys Asp Ala Arg |      |      |      |
| 1315                                                            | 1320 |      | 1325 |
| gga tct gag tgt aca cac tcc cta gct tgc tac ttc acg gtc cct gag |      |      | 4032 |
| Gly Ser Glu Cys Thr His Ser Leu Ala Cys Tyr Phe Thr Val Pro Glu |      |      |      |
| 1330                                                            | 1335 |      | 1340 |
| gtc cct cct gat gtg gcc agc atc ctg gcc agt ggc aac tgg acg cca |      |      | 4080 |
| Val Pro Pro Asp Val Ala Ser Ile Leu Ala Ser Gly Asn Trp Thr Pro |      |      |      |
| 1345                                                            | 1350 |      | 1355 |
| gaa tct cca tcc cca gct tgc caa tgc agt cag cct gga gcc cgc cgc |      |      | 4128 |
| Glu Ser Pro Ser Pro Ala Cys Gln Cys Ser Gln Pro Gly Ala Arg Arg |      |      |      |
| 1365                                                            | 1370 |      | 1375 |
| ctg ttg cca gat tgc ccg gct gga gct ggg ggt cca cca ccc ccc cag |      |      | 4176 |
| Leu Leu Pro Asp Cys Pro Ala Gly Ala Gly Gly Pro Pro Pro Pro Gln |      |      |      |
| 1380                                                            | 1385 |      | 1390 |
| gct gtg gct ggc ttg ggg gag gtg gtc cag aac ctc act ggc cga aat |      |      | 4224 |
| Ala Val Ala Gly Leu Gly Glu Val Val Gln Asn Leu Thr Gly Arg Asn |      |      |      |
| 1395                                                            | 1400 |      | 1405 |
| gtg tct gac ttt ttg gtg aag aca tac ccc agc ctg gtg cgc cga ggc |      |      | 4272 |
| Val Ser Asp Phe Leu Val Lys Thr Tyr Pro Ser Leu Val Arg Arg Gly |      |      |      |
| 1410                                                            | 1415 |      | 1420 |
| cta aag acc aag aag tgg gtg gat gag gtc aga tat ggg ggc ttc tcc |      |      | 4320 |
| Leu Lys Thr Lys Lys Trp Val Asp Glu Val Arg Tyr Gly Gly Phe Ser |      |      |      |
| 1425                                                            | 1430 |      | 1435 |
| ctg gga ggc cga gat cca gac ctg ccc aca ggg cat gag gtg gtc cgc |      |      | 4368 |
| Leu Gly Gly Arg Asp Pro Asp Leu Pro Thr Gly His Glu Val Val Arg |      |      |      |
| 1445                                                            | 1450 |      | 1455 |
| aca ttg gca gag att cgg gca ctg ctg agc ccc caa cct ggg aat gcg |      |      | 4416 |
| Thr Leu Ala Glu Ile Arg Ala Leu Leu Ser Pro Gln Pro Gly Asn Ala |      |      |      |
| 1460                                                            | 1465 |      | 1470 |
| cta gac cgt atc ctg aac aac ctc act cag tgg gcc ctt ggc ctt gat |      |      | 4464 |
| Leu Asp Arg Ile Leu Asn Asn Leu Thr Gln Trp Ala Leu Gly Leu Asp |      |      |      |
| 1475                                                            | 1480 |      | 1485 |

|                                                                 |      |
|-----------------------------------------------------------------|------|
| gct cgg aac agc ctc aag atc tgg ttc aac aac aag ggc tgg cat gcc | 4512 |
| Ala Arg Asn Ser Leu Lys Ile Trp Phe Asn Asn Lys Gly Trp His Ala |      |
| 1490 1495 1500                                                  |      |
| atg gtg gcc ttt gtg aac cga gcc aac aat gga ctc cta cat gcc ctc | 4560 |
| Met Val Ala Phe Val Asn Arg Ala Asn Asn Gly Leu Leu His Ala Leu |      |
| 1505 1510 1515 1520                                             |      |
| cta cca tct ggt cca gtc cgc cat gcc cac agc atc act aca ctc aac | 4608 |
| Leu Pro Ser Gly Pro Val Arg His Ala His Ser Ile Thr Thr Leu Asn |      |
| 1525 1530 1535                                                  |      |
| cat cct ttg aac ttg acc aag gag cag cta tct gaa gct aca ctg ata | 4656 |
| His Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ala Thr Leu Ile |      |
| 1540 1545 1550                                                  |      |
| gcc tcc tct gtg gat gtc ctt gtc tcc atc tgt gtg gtc ttc gcc atg | 4704 |
| Ala Ser Ser Val Asp Val Leu Val Ser Ile Cys Val Val Phe Ala Met |      |
| 1555 1560 1565                                                  |      |
| tca ttt gtc cca gcc agc ttt acc ctg gtc ctc ata gag gaa cgc atc | 4752 |
| Ser Phe Val Pro Ala Ser Phe Thr Leu Val Leu Ile Glu Glu Arg Ile |      |
| 1570 1575 1580                                                  |      |
| acc aga gcc aag cat ctg cag ctg gtc agc ggc ctg ccc caa acc ctc | 4800 |
| Thr Arg Ala Lys His Leu Gln Leu Val Ser Gly Leu Pro Gln Thr Leu |      |
| 1585 1590 1595 1600                                             |      |
| tat tgg ctt ggc aac ttc ctc tgg gac atg tgt aac tac ttg gtg gca | 4848 |
| Tyr Trp Leu Gly Asn Phe Leu Trp Asp Met Cys Asn Tyr Leu Val Ala |      |
| 1605 1610 1615                                                  |      |
| gtg tgc ata gtg gtg ttc atc ttc cta gcc ttt cag cag aga gcc tat | 4896 |
| Val Cys Ile Val Val Phe Ile Phe Leu Ala Phe Gln Gln Arg Ala Tyr |      |
| 1620 1625 1630                                                  |      |
| gtg gcc cca gag aac ctg cct gct ctc tta ctc ttg ctt ctg ctg tat | 4944 |
| Val Ala Pro Glu Asn Leu Pro Ala Leu Leu Leu Leu Leu Leu Tyr     |      |
| 1635 1640 1645                                                  |      |
| ggg tgg tct atc aca cca ctc atg tac cca gcc tcc ttc ttc ttc tca | 4992 |
| Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala Ser Phe Phe Phe Ser |      |
| 1650 1655 1660                                                  |      |
| gtg ccc agc acg gcc tat gtg gtg ctc acc tgc atc aac ctc ttc att | 5040 |
| Val Pro Ser Thr Ala Tyr Val Val Leu Thr Cys Ile Asn Leu Phe Ile |      |
| 1665 1670 1675 1680                                             |      |
| ggc atc aat agc agc atg gcc acc ttc gtg cta gaa ctg ctt tca gat | 5088 |
| Gly Ile Asn Ser Ser Met Ala Thr Phe Val Leu Glu Leu Leu Ser Asp |      |
| 1685 1690 1695                                                  |      |
| cag aac ctg caa gaa gtg agc cgg atc ctg aaa caa gtg ttt ctt att | 5136 |
| Gln Asn Leu Gln Glu Val Ser Arg Ile Leu Lys Gln Val Phe Leu Ile |      |
| 1700 1705 1710                                                  |      |
| ttc ccc cac ttt tgc ctt ggc cga ggg ctc att gac atg gtt cgg aac | 5184 |
| Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Met Val Arg Asn |      |

| 1715                                                            | 1720 | 1725 |      |
|-----------------------------------------------------------------|------|------|------|
| cag gcc atg gca gat gcc ttt gag cgc tta gga gac aag caa ttt cag |      |      | 5232 |
| Gln Ala Met Ala Asp Ala Phe Glu Arg Leu Gly Asp Lys Gln Phe Gln |      |      |      |
| 1730                                                            | 1735 | 1740 |      |
| tca ccc cta cgc tgg gac atc att ggc aag aac ctc ctg gcc atg atg |      |      | 5280 |
| Ser Pro Leu Arg Trp Asp Ile Ile Gly Lys Asn Leu Leu Ala Met Met |      |      |      |
| 1745                                                            | 1750 | 1755 | 1760 |
| gcc cag gga cct ctg ttc ctc ctc atc aca ctc ctg ctc caa cac cgc |      |      | 5328 |
| Ala Gln Gly Pro Leu Phe Leu Leu Ile Thr Leu Leu Leu Gln His Arg |      |      |      |
|                                                                 | 1765 | 1770 | 1775 |
| aac cgt ctc ctg cca caa tca aaa cca aga ctg ctg ccg ccc ctg ggg |      |      | 5376 |
| Asn Arg Leu Leu Pro Gln Ser Lys Pro Arg Leu Leu Pro Pro Leu Gly |      |      |      |
|                                                                 | 1780 | 1785 | 1790 |
| gag gag gat gag gat gtg gct caa gag cgt gag cgg gtg acc aag ggg |      |      | 5424 |
| Glu Glu Asp Glu Asp Val Ala Gln Glu Arg Glu Arg Val Thr Lys Gly |      |      |      |
|                                                                 | 1795 | 1800 | 1805 |
| gct acc cag ggg gat gtg cta gtc ctc agg gac ttg acc aag gtt tac |      |      | 5472 |
| Ala Thr Gln Gly Asp Val Leu Val Leu Arg Asp Leu Thr Lys Val Tyr |      |      |      |
|                                                                 | 1810 | 1815 | 1820 |
| cgt ggg cag agg aac cca gct gtg gat cgc ctg tgc tta ggg atc ccc |      |      | 5520 |
| Arg Gly Gln Arg Asn Pro Ala Val Asp Arg Leu Cys Leu Gly Ile Pro |      |      |      |
| 1825                                                            | 1830 | 1835 | 1840 |
| cct ggg gag tgt ttc ggg ctg ctg ggt gtc aac ggg gca ggg aag aca |      |      | 5568 |
| Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Thr |      |      |      |
|                                                                 | 1845 | 1850 | 1855 |
| tcc acc ttc cgc atg gtg aca ggg gac aca ctg ccc agc agt ggt gaa |      |      | 5616 |
| Ser Thr Phe Arg Met Val Thr Gly Asp Thr Leu Pro Ser Ser Gly Glu |      |      |      |
|                                                                 | 1860 | 1865 | 1870 |
| gca gta ctg gca ggc cac aac gtg gcc cag gag cgg tct gcc gca cac |      |      | 5664 |
| Ala Val Leu Ala Gly His Asn Val Ala Gln Glu Arg Ser Ala Ala His |      |      |      |
|                                                                 | 1875 | 1880 | 1885 |
| cgc agc atg ggc tac tgt ccc cag tct gat gcc atc ttc gac ctg ctg |      |      | 5712 |
| Arg Ser Met Gly Tyr Cys Pro Gln Ser Asp Ala Ile Phe Asp Leu Leu |      |      |      |
| 1890                                                            | 1895 | 1900 |      |
| acc ggc cgg gaa cat ctg gaa ctg ttt gct cgc ctg cgc ggg gtg ccc |      |      | 5760 |
| Thr Gly Arg Glu His Leu Glu Leu Phe Ala Arg Leu Arg Gly Val Pro |      |      |      |
| 1905                                                            | 1910 | 1915 | 1920 |
| gag gcc caa gtt gcc cag act gcg ctc tct ggc ctg gtg cgc ctg ggc |      |      | 5808 |
| Glu Ala Gln Val Ala Gln Thr Ala Leu Ser Gly Leu Val Arg Leu Gly |      |      |      |
|                                                                 | 1925 | 1930 | 1935 |
| ctt cct agc tat gca gac cga ccc gcg ggt acc tac agc gga ggc aac |      |      | 5856 |
| Leu Pro Ser Tyr Ala Asp Arg Pro Ala Gly Thr Tyr Ser Gly Gly Asn |      |      |      |
|                                                                 | 1940 | 1945 | 1950 |

|                                                                   |      |
|-------------------------------------------------------------------|------|
| aaa cgg aag ctg gcg aca gcc tta gct ctg gtt ggt gac cca gct gtg   | 5904 |
| Lys Arg Lys Leu Ala Thr Ala Leu Ala Leu Val Gly Asp Pro Ala Val   |      |
| 1955 1960 1965                                                    |      |
| gtc ttt ctg gac gag ccc acc aca ggc atg gac cca agt gcg cgg cga   | 5952 |
| Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Ser Ala Arg Arg   |      |
| 1970 1975 1980                                                    |      |
| ttt ctt tgg aac agc ttg ctg tcc gtg gtg cgc gag ggc cgc tcc gta   | 6000 |
| Phe Leu Trp Asn Ser Leu Leu Ser Val Val Arg Glu Gly Arg Ser Val   |      |
| 1985 1990 1995 2000                                               |      |
| gtg ctc acg tcg cac agc atg gag gag tgc gaa gcg ctc tgc acg cgc   | 6048 |
| Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg   |      |
| 2005 2010 2015                                                    |      |
| ctg gcc atc atg gtg aac ggg cgg ttc cgc tgt ctg gga agc tct cag   | 6096 |
| Leu Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu Gly Ser Ser Gln   |      |
| 2020 2025 2030                                                    |      |
| cat ctc aaa ggc agg ttc ggg gct ggc cac aca ctg act cta agg gtc   | 6144 |
| His Leu Lys Gly Arg Phe Gly Ala Gly His Thr Leu Thr Leu Arg Val   |      |
| 2035 2040 2045                                                    |      |
| cca ccg gac cag cct gag ccg gcg ata gcc ttc atc agg atc aca ttc   | 6192 |
| Pro Pro Asp Gln Pro Glu Pro Ala Ile Ala Phe Ile Arg Ile Thr Phe   |      |
| 2050 2055 2060                                                    |      |
| cct ggg gct gaa ctc cgg gag gtg cac ggc agc cgt ctg cgc ttc caa   | 6240 |
| Pro Gly Ala Glu Leu Arg Glu Val His Gly Ser Arg Leu Arg Phe Gln   |      |
| 2065 2070 2075 2080                                               |      |
| ctg cca ccg ggg ggc aga tgc acc ctg aca cga gtg ttc agg gag ctg   | 6288 |
| Leu Pro Pro Gly Gly Arg Cys Thr Leu Thr Arg Val Phe Arg Glu Leu   |      |
| 2085 2090 2095                                                    |      |
| gct gcc cag ggc agg gcc cac ggt gtg gag gac ttc tct gtg agc cag   | 6336 |
| Ala Ala Gln Gly Arg Ala His Gly Val Glu Asp Phe Ser Val Ser Gln   |      |
| 2100 2105 2110                                                    |      |
| acc act ctg gag gag gtg ttc cta tat ttc tcc aaa gac caa ggg gaa   | 6384 |
| Thr Thr Leu Glu Glu Val Phe Leu Tyr Phe Ser Lys Asp Gln Gly Glu   |      |
| 2115 2120 2125                                                    |      |
| gag gaa gag agc agt cgg cag gag gct gaa gaa gag gag gtt tcc aaa   | 6432 |
| Glu Glu Glu Ser Ser Arg Gln Glu Ala Glu Glu Glu Glu Val Ser Lys   |      |
| 2130 2135 2140                                                    |      |
| cct ggc cgg cag cat ccc aaa cgt gtc agc cga ttc ctg gaa gac ccc   | 6480 |
| Pro Gly Arg Gln His Pro Lys Arg Val Ser Arg Phe Leu Glu Asp Pro   |      |
| 2145 2150 2155 2160                                               |      |
| agc tct gtg gag acc atg atc tga gcatgcctgc cttgggactg agtggcaaag  | 6534 |
| Ser Ser Val Glu Thr Met Ile                                       |      |
| 2165                                                              |      |
| ctcagacaga ggatctctgt accatacgct ggctcccaga aagccttggg ctctggggga | 6594 |

aataaaaaaga aactagaatg agaaaaaaaaa aaaaaaaaaa

6633

&lt;210&gt; 2

&lt;211&gt; 2167

&lt;212&gt; PRT

&lt;213&gt; Mus musculus

&lt;400&gt; 2

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ala | Phe | Cys | Thr | Gln | Leu | Met | Leu | Leu | Leu | Trp | Lys | Asn | Tyr | Thr |
| 1   |     |     |     | 5   |     |     |     | 10  |     |     |     |     |     | 15  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Tyr | Arg | Arg | Arg | Gln | Pro | Ile | Gln | Leu | Leu | Val | Glu | Leu | Leu | Trp | Pro |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Phe | Leu | Phe | Phe | Ile | Leu | Val | Ala | Val | Arg | His | Ser | His | Pro | Pro |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Glu | His | His | Glu | Cys | His | Phe | Pro | Asn | Lys | Pro | Leu | Pro | Ser | Ala |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Thr | Val | Pro | Trp | Leu | Gln | Gly | Leu | Val | Cys | Asn | Val | Asn | Asn | Ser |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Cys | Phe | Gln | His | Pro | Thr | Pro | Gly | Glu | Lys | Pro | Gly | Val | Leu | Ser | Asn |
|     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Phe | Lys | Asp | Ser | Leu | Ile | Ser | Arg | Leu | Leu | Ala | Asp | Thr | Arg | Thr | Val |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Gly | Gly | His | Ser | Ile | Gln | Asp | Met | Leu | Asp | Ala | Leu | Gly | Lys | Leu |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ile | Pro | Val | Leu | Arg | Ala | Val | Gly | Gly | Gly | Ala | Arg | Pro | Gln | Glu | Ser |
|     | 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Gln | Pro | Thr | Ser | Gln | Gly | Ser | Val | Thr | Lys | Leu | Leu | Glu | Lys | Ile |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Gln | Arg | Ala | Ser | Leu | Asp | Pro | Val | Leu | Gly | Gln | Ala | Gln | Asp | Ser |
|     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Arg | Lys | Phe | Ser | Asp | Ala | Ile | Arg | Asp | Leu | Ala | Gln | Glu | Leu | Leu |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     | 190 |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Leu | Pro | Ser | Leu | Met | Glu | Leu | Arg | Ala | Leu | Leu | Arg | Arg | Pro | Arg |
|     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Ser | Ala | Gly | Ser | Leu | Glu | Leu | Val | Ser | Glu | Ala | Leu | Cys | Ser | Thr |
|     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lys | Gly | Pro | Ser | Ser | Pro | Gly | Gly | Leu | Ser | Leu | Asn | Trp | Tyr | Glu | Ala |
| 225 |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Gln | Leu | Asn | Glu | Phe | Met | Gly | Pro | Glu | Val | Ala | Pro | Ala | Leu | Pro |
|     |     |     | 245 |     |     |     |     |     | 250 |     |     |     |     | 255 |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Asn | Ser | Leu | Ser | Pro | Ala | Cys | Ser | Glu | Phe | Val | Gly | Thr | Leu | Asp |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

| 260 |     |     |     |     |     |     |     |     |     | 265 |     |     |     |     | 270 |  |  |  |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|--|
| Asp | His | Pro | Val | Ser | Arg | Leu | Leu | Trp | Arg | Arg | Leu | Lys | Pro | Leu | Ile |  |  |  |  |
|     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |  |  |  |  |
| Leu | Gly | Lys | Ile | Leu | Phe | Ala | Pro | Asp | Thr | Asn | Phe | Thr | Arg | Lys | Leu |  |  |  |  |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |  |  |  |  |
| Met | Ala | Gln | Val | Asn | Gln | Thr | Phe | Glu | Glu | Leu | Ala | Leu | Leu | Arg | Asp |  |  |  |  |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |  |  |  |  |
| Leu | His | Glu | Leu | Trp | Gly | Val | Leu | Gly | Pro | Gln | Ile | Phe | Asn | Phe | Met |  |  |  |  |
|     |     |     |     | 325 |     |     |     | 330 |     |     |     |     |     | 335 |     |  |  |  |  |
| Asn | Asp | Ser | Thr | Asn | Val | Ala | Met | Leu | Gln | Arg | Leu | Leu | Asp | Val | Gly |  |  |  |  |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |     |     |  |  |  |  |
| Gly | Thr | Gly | Gln | Arg | Gln | Gln | Thr | Pro | Arg | Ala | Gln | Lys | Lys | Leu | Glu |  |  |  |  |
|     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |  |  |  |  |
| Ala | Ile | Lys | Asp | Phe | Leu | Asp | Pro | Ser | Arg | Gly | Gly | Tyr | Ser | Trp | Arg |  |  |  |  |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |  |  |  |  |
| Glu | Ala | His | Ala | Asp | Met | Gly | Arg | Leu | Ala | Gly | Ile | Leu | Gly | Gln | Met |  |  |  |  |
| 385 |     |     |     |     | 390 |     |     |     |     | 395 |     |     |     |     | 400 |  |  |  |  |
| Met | Glu | Cys | Val | Ser | Leu | Asp | Lys | Leu | Glu | Ala | Val | Pro | Ser | Glu | Glu |  |  |  |  |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |     |  |  |  |  |
| Ala | Leu | Val | Ser | Arg | Ala | Leu | Glu | Leu | Leu | Gly | Glu | Arg | Arg | Leu | Trp |  |  |  |  |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     |     | 430 |     |  |  |  |  |
| Ala | Gly | Ile | Val | Phe | Leu | Ser | Pro | Glu | His | Pro | Leu | Asp | Pro | Ser | Glu |  |  |  |  |
|     |     | 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |  |  |  |  |
| Leu | Ser | Ser | Pro | Ala | Leu | Ser | Pro | Gly | His | Leu | Arg | Phe | Lys | Ile | Arg |  |  |  |  |
|     | 450 |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |  |  |  |  |
| Met | Asp | Ile | Asp | Asp | Val | Thr | Arg | Thr | Asn | Lys | Ile | Arg | Asp | Lys | Phe |  |  |  |  |
| 465 |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |  |  |  |  |
| Trp | Asp | Pro | Gly | Pro | Ser | Ala | Asp | Pro | Phe | Met | Asp | Leu | Arg | Tyr | Val |  |  |  |  |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     | 495 |     |  |  |  |  |
| Trp | Gly | Gly | Phe | Val | Tyr | Leu | Gln | Asp | Leu | Leu | Glu | Gln | Ala | Ala | Val |  |  |  |  |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |     |     |  |  |  |  |
| Arg | Val | Leu | Gly | Gly | Gly | Asn | Ser | Arg | Thr | Gly | Leu | Tyr | Leu | Gln | Gln |  |  |  |  |
|     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |     |  |  |  |  |
| Met | Pro | His | Pro | Cys | Tyr | Val | Asp | Asp | Val | Phe | Leu | Arg | Val | Leu | Ser |  |  |  |  |
|     | 530 |     |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |  |  |  |  |
| Arg | Ser | Leu | Pro | Leu | Phe | Leu | Thr | Leu | Ala | Trp | Ile | Tyr | Ser | Val | Ala |  |  |  |  |
| 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |  |  |  |  |
| Leu | Thr | Val | Lys | Ala | Val | Val | Arg | Glu | Lys | Glu | Thr | Arg | Leu | Arg | Glu |  |  |  |  |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |  |  |  |  |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Met | Arg | Ala | Met | Gly | Leu | Ser | Arg | Ala | Val | Leu | Trp | Leu | Gly | Trp | 580 | 585 | 590 |
| Phe | Leu | Ser | Cys | Leu | Gly | Pro | Phe | Leu | Val | Ser | Ala | Ala | Leu | Leu | Val | 595 | 600 | 605 |
| Leu | Val | Leu | Lys | Leu | Gly | Asn | Ile | Leu | Pro | Tyr | Ser | His | Pro | Val | Val | 610 | 615 | 620 |
| Ile | Phe | Leu | Phe | Leu | Ala | Ala | Phe | Ala | Val | Ala | Thr | Val | Ala | Gln | Ser | 625 | 630 | 635 |
| Phe | Leu | Leu | Ser | Ala | Phe | Phe | Ser | Arg | Ala | Asn | Leu | Ala | Ala | Ala | Cys | 645 | 650 | 655 |
| Gly | Gly | Leu | Ala | Tyr | Phe | Ala | Leu | Tyr | Leu | Pro | Tyr | Val | Leu | Cys | Val | 660 | 665 | 670 |
| Ala | Trp | Arg | Glu | Arg | Leu | His | Leu | Gly | Gly | Leu | Leu | Ala | Ala | Ser | Leu | 675 | 680 | 685 |
| Leu | Ser | Pro | Val | Ala | Phe | Gly | Phe | Gly | Cys | Glu | Ser | Leu | Ala | Leu | Leu | 690 | 695 | 700 |
| Glu | Glu | Gln | Gly | Asp | Gly | Ala | Gln | Trp | His | Asn | Leu | Gly | Thr | Gly | Pro | 705 | 710 | 715 |
| Ala | Glu | Asp | Val | Phe | Ser | Leu | Ala | Gln | Val | Ser | Ala | Phe | Leu | Leu | Leu | 725 | 730 | 735 |
| Asp | Ala | Val | Ile | Tyr | Gly | Leu | Ala | Leu | Trp | Tyr | Leu | Glu | Ala | Val | Cys | 740 | 745 | 750 |
| Pro | Gly | Gln | Tyr | Gly | Ile | Pro | Glu | Pro | Trp | Asn | Phe | Pro | Phe | Arg | Arg | 755 | 760 | 765 |
| Ser | Tyr | Trp | Cys | Gly | Pro | Gly | Pro | Pro | Lys | Ser | Ser | Val | Leu | Ala | Pro | 770 | 775 | 780 |
| Ala | Pro | Gln | Asp | Pro | Lys | Val | Leu | Val | Glu | Glu | Pro | Pro | Leu | Gly | Leu | 785 | 790 | 795 |
| Val | Pro | Gly | Val | Ser | Ile | Arg | Gly | Leu | Lys | Lys | His | Phe | Arg | Gly | Cys | 805 | 810 | 815 |
| Pro | Gln | Pro | Ala | Leu | Gln | Gly | Leu | Asn | Leu | Asp | Phe | Tyr | Glu | Gly | His | 820 | 825 | 830 |
| Ile | Thr | Ala | Phe | Leu | Gly | His | Asn | Gly | Ala | Gly | Lys | Thr | Thr | Thr | Leu | 835 | 840 | 845 |
| Ser | Ile | Leu | Ser | Gly | Leu | Phe | Pro | Pro | Ser | Ser | Gly | Ser | Ala | Ser | Ile | 850 | 855 | 860 |
| Leu | Gly | His | Asp | Val | Gln | Thr | Asn | Met | Ala | Ala | Ile | Arg | Pro | His | Leu | 865 | 870 | 875 |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 880 |     |     |

Gly Ile Cys Pro Gln Tyr Asn Val Leu Phe Asp Met Leu Thr Val Glu  
 885 890 895  
 Glu His Val Trp Phe Tyr Gly Arg Leu Lys Gly Val Ser Ala Ala Ala  
 900 905 910  
 Met Gly Pro Glu Arg Glu Arg Leu Ile Arg Asp Val Gly Leu Thr Leu  
 915 920 925  
 Lys Arg Asp Thr Gln Thr Arg His Leu Ser Gly Gly Met Gln Arg Lys  
 930 935 940  
 Leu Ser Val Ala Ile Ala Phe Val Gly Gly Ser Arg Val Val Ile Met  
 945 950 955 960  
 Asp Glu Pro Thr Ala Gly Val Asp Pro Ala Ser Arg Arg Gly Ile Trp  
 965 970 975  
 Glu Leu Leu Leu Lys Tyr Arg Glu Gly Arg Thr Leu Ile Leu Ser Thr  
 980 985 990  
 His His Leu Asp Glu Ala Glu Leu Leu Gly Asp Arg Val Ala Met Val  
 995 1000 1005  
 Ala Gly Gly Ser Leu Cys Cys Cys Gly Ser Pro Leu Phe Leu Arg Arg  
 1010 1015 1020  
 His Leu Gly Cys Gly Tyr Tyr Leu Thr Leu Val Lys Ser Ser Gln Ser  
 1025 1030 1035 1040  
 Leu Val Thr His Asp Ala Lys Gly Asp Ser Glu Asp Pro Arg Arg Glu  
 1045 1050 1055  
 Lys Lys Ser Asp Gly Asn Gly Arg Thr Ser Asp Thr Ala Phe Thr Arg  
 1060 1065 1070  
 Gly Thr Ser Asp Lys Ser Asn Gln Ala Pro Ala Pro Gly Ala Val Pro  
 1075 1080 1085  
 Ile Thr Pro Ser Thr Ala Arg Ile Leu Glu Leu Val Gln Gln His Val  
 1090 1095 1100  
 Pro Gly Ala Gln Leu Val Glu Asp Leu Pro His Glu Leu Leu Leu Val  
 1105 1110 1115 1120  
 Leu Pro Tyr Ala Gly Ala Leu Asp Gly Ser Phe Ala Met Val Phe Gln  
 1125 1130 1135  
 Glu Leu Asp Gln Gln Leu Glu Leu Leu Gly Leu Thr Gly Tyr Gly Ile  
 1140 1145 1150  
 Ser Asp Thr Asn Leu Glu Glu Ile Phe Leu Lys Val Val Glu Asp Ala  
 1155 1160 1165  
 His Arg Glu Gly Gly Asp Ser Arg Pro Gln Leu His Leu Arg Thr Cys  
 1170 1175 1180  
 Thr Pro Gln Pro Pro Thr Gly Pro Glu Ala Ser Val Leu Glu Asn Gly

|                                                                 |      |      |      |
|-----------------------------------------------------------------|------|------|------|
| 1185                                                            | 1190 | 1195 | 1200 |
| Glu Leu Ala Lys Leu Val Leu Asp Pro Gln Ala Pro Gln Gly Leu Ala | 1205 | 1210 | 1215 |
| Pro Asn Ala Ala Gln Val Gln Gly Trp Thr Leu Thr Cys Gln Gln Leu | 1220 | 1225 | 1230 |
| Arg Ala Leu Leu His Lys Arg Phe Leu Leu Ala Arg Arg Ser Arg Arg | 1235 | 1240 | 1245 |
| Gly Leu Phe Ala Gln Val Val Leu Pro Ala Leu Phe Val Gly Leu Ala | 1250 | 1255 | 1260 |
| Leu Phe Phe Ser Leu Ile Val Pro Pro Phe Gly Gln Tyr Pro Pro Leu | 1265 | 1270 | 1275 |
| Gln Leu Ser Pro Ala Met Tyr Gly Pro Gln Val Ser Phe Phe Ser Glu | 1285 | 1290 | 1295 |
| Asp Ala Pro Gly Asp Pro Asn Arg Met Lys Leu Leu Glu Ala Leu Leu | 1300 | 1305 | 1310 |
| Gly Glu Ala Gly Leu Gln Glu Pro Ser Met Gln Asp Lys Asp Ala Arg | 1315 | 1320 | 1325 |
| Gly Ser Glu Cys Thr His Ser Leu Ala Cys Tyr Phe Thr Val Pro Glu | 1330 | 1335 | 1340 |
| Val Pro Pro Asp Val Ala Ser Ile Leu Ala Ser Gly Asn Trp Thr Pro | 1345 | 1350 | 1355 |
| Glu Ser Pro Ser Pro Ala Cys Gln Cys Ser Gln Pro Gly Ala Arg Arg | 1365 | 1370 | 1375 |
| Leu Leu Pro Asp Cys Pro Ala Gly Ala Gly Gly Pro Pro Pro Pro Gln | 1380 | 1385 | 1390 |
| Ala Val Ala Gly Leu Gly Glu Val Val Gln Asn Leu Thr Gly Arg Asn | 1395 | 1400 | 1405 |
| Val Ser Asp Phe Leu Val Lys Thr Tyr Pro Ser Leu Val Arg Arg Gly | 1410 | 1415 | 1420 |
| Leu Lys Thr Lys Lys Trp Val Asp Glu Val Arg Tyr Gly Gly Phe Ser | 1425 | 1430 | 1435 |
| Leu Gly Gly Arg Asp Pro Asp Leu Pro Thr Gly His Glu Val Val Arg | 1445 | 1450 | 1455 |
| Thr Leu Ala Glu Ile Arg Ala Leu Leu Ser Pro Gln Pro Gly Asn Ala | 1460 | 1465 | 1470 |
| Leu Asp Arg Ile Leu Asn Asn Leu Thr Gln Trp Ala Leu Gly Leu Asp | 1475 | 1480 | 1485 |
| Ala Arg Asn Ser Leu Lys Ile Trp Phe Asn Asn Lys Gly Trp His Ala | 1490 | 1495 | 1500 |

Met Val Ala Phe Val Asn Arg Ala Asn Asn Gly Leu Leu His Ala Leu  
 1505 1510 1515 1520  
 Leu Pro Ser Gly Pro Val Arg His Ala His Ser Ile Thr Thr Leu Asn  
 1525 1530 1535  
 His Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ala Thr Leu Ile  
 1540 1545 1550  
 Ala Ser Ser Val Asp Val Leu Val Ser Ile Cys Val Val Phe Ala Met  
 1555 1560 1565  
 Ser Phe Val Pro Ala Ser Phe Thr Leu Val Leu Ile Glu Glu Arg Ile  
 1570 1575 1580  
 Thr Arg Ala Lys His Leu Gln Leu Val Ser Gly Leu Pro Gln Thr Leu  
 1585 1590 1595 1600  
 Tyr Trp Leu Gly Asn Phe Leu Trp Asp Met Cys Asn Tyr Leu Val Ala  
 1605 1610 1615  
 Val Cys Ile Val Val Phe Ile Phe Leu Ala Phe Gln Gln Arg Ala Tyr  
 1620 1625 1630  
 Val Ala Pro Glu Asn Leu Pro Ala Leu Leu Leu Leu Leu Leu Tyr  
 1635 1640 1645  
 Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala Ser Phe Phe Phe Ser  
 1650 1655 1660  
 Val Pro Ser Thr Ala Tyr Val Val Leu Thr Cys Ile Asn Leu Phe Ile  
 1665 1670 1675 1680  
 Gly Ile Asn Ser Ser Met Ala Thr Phe Val Leu Glu Leu Leu Ser Asp  
 1685 1690 1695  
 Gln Asn Leu Gln Glu Val Ser Arg Ile Leu Lys Gln Val Phe Leu Ile  
 1700 1705 1710  
 Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Met Val Arg Asn  
 1715 1720 1725  
 Gln Ala Met Ala Asp Ala Phe Glu Arg Leu Gly Asp Lys Gln Phe Gln  
 1730 1735 1740  
 Ser Pro Leu Arg Trp Asp Ile Ile Gly Lys Asn Leu Leu Ala Met Met  
 1745 1750 1755 1760  
 Ala Gln Gly Pro Leu Phe Leu Leu Ile Thr Leu Leu Leu Gln His Arg  
 1765 1770 1775  
 Asn Arg Leu Leu Pro Gln Ser Lys Pro Arg Leu Leu Pro Pro Leu Gly  
 1780 1785 1790  
 Glu Glu Asp Glu Asp Val Ala Gln Glu Arg Glu Arg Val Thr Lys Gly  
 1795 1800 1805

Ala Thr Gln Gly Asp Val Leu Val Leu Arg Asp Leu Thr Lys Val Tyr  
 1810 1815 1820  
 Arg Gly Gln Arg Asn Pro Ala Val Asp Arg Leu Cys Leu Gly Ile Pro  
 1825 1830 1835 1840  
 Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Thr  
 1845 1850 1855  
 Ser Thr Phe Arg Met Val Thr Gly Asp Thr Leu Pro Ser Ser Gly Glu  
 1860 1865 1870  
 Ala Val Leu Ala Gly His Asn Val Ala Gln Glu Arg Ser Ala Ala His  
 1875 1880 1885  
 Arg Ser Met Gly Tyr Cys Pro Gln Ser Asp Ala Ile Phe Asp Leu Leu  
 1890 1895 1900  
 Thr Gly Arg Glu His Leu Glu Leu Phe Ala Arg Leu Arg Gly Val Pro  
 1905 1910 1915 1920  
 Glu Ala Gln Val Ala Gln Thr Ala Leu Ser Gly Leu Val Arg Leu Gly  
 1925 1930 1935  
 Leu Pro Ser Tyr Ala Asp Arg Pro Ala Gly Thr Tyr Ser Gly Gly Asn  
 1940 1945 1950  
 Lys Arg Lys Leu Ala Thr Ala Leu Ala Leu Val Gly Asp Pro Ala Val  
 1955 1960 1965  
 Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Ser Ala Arg Arg  
 1970 1975 1980  
 Phe Leu Trp Asn Ser Leu Leu Ser Val Val Arg Glu Gly Arg Ser Val  
 1985 1990 1995 2000  
 Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg  
 2005 2010 2015  
 Leu Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu Gly Ser Ser Gln  
 2020 2025 2030  
 His Leu Lys Gly Arg Phe Gly Ala Gly His Thr Leu Thr Leu Arg Val  
 2035 2040 2045  
 Pro Pro Asp Gln Pro Glu Pro Ala Ile Ala Phe Ile Arg Ile Thr Phe  
 2050 2055 2060  
 Pro Gly Ala Glu Leu Arg Glu Val His Gly Ser Arg Leu Arg Phe Gln  
 2065 2070 2075 2080  
 Leu Pro Pro Gly Gly Arg Cys Thr Leu Thr Arg Val Phe Arg Glu Leu  
 2085 2090 2095  
 Ala Ala Gln Gly Arg Ala His Gly Val Glu Asp Phe Ser Val Ser Gln  
 2100 2105 2110  
 Thr Thr Leu Glu Glu Val Phe Leu Tyr Phe Ser Lys Asp Gln Gly Glu

2115                      2120                      2125  
 Glu Glu Glu Ser Ser Arg Gln Glu Ala Glu Glu Glu Glu Val Ser Lys  
     2130                      2135                      2140  
 Pro Gly Arg Gln His Pro Lys Arg Val Ser Arg Phe Leu Glu Asp Pro  
     2145                      2150                      2155                      2160  
 Ser Ser Val Glu Thr Met Ile  
                                  2165  
  
 <210> 3  
 <211> 2121  
 <212> PRT  
 <213> Mus musculus  
  
 <400> 3  
 Pro Pro Leu Glu His His Glu Cys His Phe Pro Asn Lys Pro Leu Pro  
     1                      5                      10                      15  
 Ser Ala Gly Thr Val Pro Trp Leu Gln Gly Leu Val Cys Asn Val Asn  
                                  20                      25                      30  
 Asn Ser Cys Phe Gln His Pro Thr Pro Gly Glu Lys Pro Gly Val Leu  
                                  35                      40                      45  
 Ser Asn Phe Lys Asp Ser Leu Ile Ser Arg Leu Leu Ala Asp Thr Arg  
                                  50                      55                      60  
 Thr Val Leu Gly Gly His Ser Ile Gln Asp Met Leu Asp Ala Leu Gly  
     65                      70                      75                      80  
 Lys Leu Ile Pro Val Leu Arg Ala Val Gly Gly Gly Ala Arg Pro Gln  
                                  85                      90                      95  
 Glu Ser Asp Gln Pro Thr Ser Gln Gly Ser Val Thr Lys Leu Leu Glu  
                                  100                      105                      110  
 Lys Ile Leu Gln Arg Ala Ser Leu Asp Pro Val Leu Gly Gln Ala Gln  
                                  115                      120                      125  
 Asp Ser Met Arg Lys Phe Ser Asp Ala Ile Arg Asp Leu Ala Gln Glu  
     130                      135                      140  
 Leu Leu Thr Leu Pro Ser Leu Met Glu Leu Arg Ala Leu Leu Arg Arg  
     145                      150                      155                      160  
 Pro Arg Gly Ser Ala Gly Ser Leu Glu Leu Val Ser Glu Ala Leu Cys  
                                  165                      170                      175  
 Ser Thr Lys Gly Pro Ser Ser Pro Gly Gly Leu Ser Leu Asn Trp Tyr  
                                  180                      185                      190  
 Glu Ala Asn Gln Leu Asn Glu Phe Met Gly Pro Glu Val Ala Pro Ala  
                                  195                      200                      205  
 Leu Pro Asp Asn Ser Leu Ser Pro Ala Cys Ser Glu Phe Val Gly Thr

| 210                                                                                | 215 | 220 |
|------------------------------------------------------------------------------------|-----|-----|
| Leu Asp Asp His Pro Val Ser Arg Leu Leu Trp Arg Arg Leu Lys Pro<br>225 230 235 240 |     |     |
| Leu Ile Leu Gly Lys Ile Leu Phe Ala Pro Asp Thr Asn Phe Thr Arg<br>245 250 255     |     |     |
| Lys Leu Met Ala Gln Val Asn Gln Thr Phe Glu Glu Leu Ala Leu Leu<br>260 265 270     |     |     |
| Arg Asp Leu His Glu Leu Trp Gly Val Leu Gly Pro Gln Ile Phe Asn<br>275 280 285     |     |     |
| Phe Met Asn Asp Ser Thr Asn Val Ala Met Leu Gln Arg Leu Leu Asp<br>290 295 300     |     |     |
| Val Gly Gly Thr Gly Gln Arg Gln Gln Thr Pro Arg Ala Gln Lys Lys<br>305 310 315 320 |     |     |
| Leu Glu Ala Ile Lys Asp Phe Leu Asp Pro Ser Arg Gly Gly Tyr Ser<br>325 330 335     |     |     |
| Trp Arg Glu Ala His Ala Asp Met Gly Arg Leu Ala Gly Ile Leu Gly<br>340 345 350     |     |     |
| Gln Met Met Glu Cys Val Ser Leu Asp Lys Leu Glu Ala Val Pro Ser<br>355 360 365     |     |     |
| Glu Glu Ala Leu Val Ser Arg Ala Leu Glu Leu Leu Gly Glu Arg Arg<br>370 375 380     |     |     |
| Leu Trp Ala Gly Ile Val Phe Leu Ser Pro Glu His Pro Leu Asp Pro<br>385 390 395 400 |     |     |
| Ser Glu Leu Ser Ser Pro Ala Leu Ser Pro Gly His Leu Arg Phe Lys<br>405 410 415     |     |     |
| Ile Arg Met Asp Ile Asp Asp Val Thr Arg Thr Asn Lys Ile Arg Asp<br>420 425 430     |     |     |
| Lys Phe Trp Asp Pro Gly Pro Ser Ala Asp Pro Phe Met Asp Leu Arg<br>435 440 445     |     |     |
| Tyr Val Trp Gly Gly Phe Val Tyr Leu Gln Asp Leu Leu Glu Gln Ala<br>450 455 460     |     |     |
| Ala Val Arg Val Leu Gly Gly Gly Asn Ser Arg Thr Gly Leu Tyr Leu<br>465 470 475 480 |     |     |
| Gln Gln Met Pro His Pro Cys Tyr Val Asp Asp Val Phe Leu Arg Val<br>485 490 495     |     |     |
| Leu Ser Arg Ser Leu Pro Leu Phe Leu Thr Leu Ala Trp Ile Tyr Ser<br>500 505 510     |     |     |
| Val Ala Leu Thr Val Lys Ala Val Val Arg Glu Lys Glu Thr Arg Leu<br>515 520 525     |     |     |

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

His Leu Gly Ile Cys Pro Gln Tyr Asn Val Leu Phe Asp Met Leu Thr  
 835 840 845  
 Val Glu Glu His Val Trp Phe Tyr Gly Arg Leu Lys Gly Val Ser Ala  
 850 855 860  
 Ala Ala Met Gly Pro Glu Arg Glu Arg Leu Ile Arg Asp Val Gly Leu  
 865 870 875 880  
 Thr Leu Lys Arg Asp Thr Gln Thr Arg His Leu Ser Gly Gly Met Gln  
 885 890 895  
 Arg Lys Leu Ser Val Ala Ile Ala Phe Val Gly Gly Ser Arg Val Val  
 900 905 910  
 Ile Met Asp Glu Pro Thr Ala Gly Val Asp Pro Ala Ser Arg Arg Gly  
 915 920 925  
 Ile Trp Glu Leu Leu Leu Lys Tyr Arg Glu Gly Arg Thr Leu Ile Leu  
 930 935 940  
 Ser Thr His His Leu Asp Glu Ala Glu Leu Leu Gly Asp Arg Val Ala  
 945 950 955 960  
 Met Val Ala Gly Gly Ser Leu Cys Cys Cys Gly Ser Pro Leu Phe Leu  
 965 970 975  
 Arg Arg His Leu Gly Cys Gly Tyr Tyr Leu Thr Leu Val Lys Ser Ser  
 980 985 990  
 Gln Ser Leu Val Thr His Asp Ala Lys Gly Asp Ser Glu Asp Pro Arg  
 995 1000 1005  
 Arg Glu Lys Lys Ser Asp Gly Asn Gly Arg Thr Ser Asp Thr Ala Phe  
 1010 1015 1020  
 Thr Arg Gly Thr Ser Asp Lys Ser Asn Gln Ala Pro Ala Pro Gly Ala  
 1025 1030 1035 1040  
 Val Pro Ile Thr Pro Ser Thr Ala Arg Ile Leu Glu Leu Val Gln Gln  
 1045 1050 1055  
 His Val Pro Gly Ala Gln Leu Val Glu Asp Leu Pro His Glu Leu Leu  
 1060 1065 1070  
 Leu Val Leu Pro Tyr Ala Gly Ala Leu Asp Gly Ser Phe Ala Met Val  
 1075 1080 1085  
 Phe Gln Glu Leu Asp Gln Gln Leu Glu Leu Leu Gly Leu Thr Gly Tyr  
 1090 1095 1100  
 Gly Ile Ser Asp Thr Asn Leu Glu Glu Ile Phe Leu Lys Val Val Glu  
 1105 1110 1115 1120  
 Asp Ala His Arg Glu Gly Gly Asp Ser Arg Pro Gln Leu His Leu Arg  
 1125 1130 1135  
 Thr Cys Thr Pro Gln Pro Pro Thr Gly Pro Glu Ala Ser Val Leu Glu

|                                     |                                 |      |
|-------------------------------------|---------------------------------|------|
| 1140                                | 1145                            | 1150 |
| Asn Gly Glu Leu Ala Lys Leu Val     | Leu Asp Pro Gln Ala Pro Gln Gly |      |
| 1155                                | 1160                            | 1165 |
| Leu Ala Pro Asn Ala Ala Gln Val     | Gln Gly Trp Thr Leu Thr Cys Gln |      |
| 1170                                | 1175                            | 1180 |
| Gln Leu Arg Ala Leu Leu His Lys Arg | Phe Leu Leu Ala Arg Arg Ser     |      |
| 1185                                | 1190                            | 1195 |
| Arg Arg Gly Leu Phe Ala Gln Val     | Val Leu Pro Ala Leu Phe Val Gly |      |
| 1205                                | 1210                            | 1215 |
| Leu Ala Leu Phe Phe Ser Leu Ile Val | Pro Pro Phe Gly Gln Tyr Pro     |      |
| 1220                                | 1225                            | 1230 |
| Pro Leu Gln Leu Ser Pro Ala Met Tyr | Gly Pro Gln Val Ser Phe Phe     |      |
| 1235                                | 1240                            | 1245 |
| Ser Glu Asp Ala Pro Gly Asp Pro Asn | Arg Met Lys Leu Leu Glu Ala     |      |
| 1250                                | 1255                            | 1260 |
| Leu Leu Gly Glu Ala Gly Leu Gln Glu | Pro Ser Met Gln Asp Lys Asp     |      |
| 1265                                | 1270                            | 1275 |
| Ala Arg Gly Ser Glu Cys Thr His Ser | Leu Ala Cys Tyr Phe Thr Val     |      |
| 1285                                | 1290                            | 1295 |
| Pro Glu Val Pro Pro Asp Val Ala Ser | Ile Leu Ala Ser Gly Asn Trp     |      |
| 1300                                | 1305                            | 1310 |
| Thr Pro Glu Ser Pro Ser Pro Ala Cys | Gln Cys Ser Gln Pro Gly Ala     |      |
| 1315                                | 1320                            | 1325 |
| Arg Arg Leu Leu Pro Asp Cys Pro Ala | Gly Ala Gly Gly Pro Pro Pro     |      |
| 1330                                | 1335                            | 1340 |
| Pro Gln Ala Val Ala Gly Leu Gly Glu | Val Val Gln Asn Leu Thr Gly     |      |
| 1345                                | 1350                            | 1355 |
| Arg Asn Val Ser Asp Phe Leu Val Lys | Thr Tyr Pro Ser Leu Val Arg     |      |
| 1365                                | 1370                            | 1375 |
| Arg Gly Leu Lys Thr Lys Lys Trp Val | Asp Glu Val Arg Tyr Gly Gly     |      |
| 1380                                | 1385                            | 1390 |
| Phe Ser Leu Gly Gly Arg Asp Pro Asp | Leu Pro Thr Gly His Glu Val     |      |
| 1395                                | 1400                            | 1405 |
| Val Arg Thr Leu Ala Glu Ile Arg Ala | Leu Leu Ser Pro Gln Pro Gly     |      |
| 1410                                | 1415                            | 1420 |
| Asn Ala Leu Asp Arg Ile Leu Asn Asn | Leu Thr Gln Trp Ala Leu Gly     |      |
| 1425                                | 1430                            | 1435 |
| Leu Asp Ala Arg Asn Ser Leu Lys Ile | Trp Phe Asn Asn Lys Gly Trp     |      |
| 1445                                | 1450                            | 1455 |

His Ala Met Val Ala Phe Val Asn Arg Ala Asn Asn Gly Leu Leu His  
 1460 1465 1470  
 Ala Leu Leu Pro Ser Gly Pro Val Arg His Ala His Ser Ile Thr Thr  
 1475 1480 1485  
 Leu Asn His Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ala Thr  
 1490 1495 1500  
 Leu Ile Ala Ser Ser Val Asp Val Leu Val Ser Ile Cys Val Val Phe  
 1505 1510 1515 1520  
 Ala Met Ser Phe Val Pro Ala Ser Phe Thr Leu Val Leu Ile Glu Glu  
 1525 1530 1535  
 Arg Ile Thr Arg Ala Lys His Leu Gln Leu Val Ser Gly Leu Pro Gln  
 1540 1545 1550  
 Thr Leu Tyr Trp Leu Gly Asn Phe Leu Trp Asp Met Cys Asn Tyr Leu  
 1555 1560 1565  
 Val Ala Val Cys Ile Val Val Phe Ile Phe Leu Ala Phe Gln Gln Arg  
 1570 1575 1580  
 Ala Tyr Val Ala Pro Glu Asn Leu Pro Ala Leu Leu Leu Leu Leu  
 1585 1590 1595 1600  
 Leu Tyr Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala Ser Phe Phe  
 1605 1610 1615  
 Phe Ser Val Pro Ser Thr Ala Tyr Val Val Leu Thr Cys Ile Asn Leu  
 1620 1625 1630  
 Phe Ile Gly Ile Asn Ser Ser Met Ala Thr Phe Val Leu Glu Leu Leu  
 1635 1640 1645  
 Ser Asp Gln Asn Leu Gln Glu Val Ser Arg Ile Leu Lys Gln Val Phe  
 1650 1655 1660  
 Leu Ile Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Met Val  
 1665 1670 1675 1680  
 Arg Asn Gln Ala Met Ala Asp Ala Phe Glu Arg Leu Gly Asp Lys Gln  
 1685 1690 1695  
 Phe Gln Ser Pro Leu Arg Trp Asp Ile Ile Gly Lys Asn Leu Leu Ala  
 1700 1705 1710  
 Met Met Ala Gln Gly Pro Leu Phe Leu Leu Ile Thr Leu Leu Leu Gln  
 1715 1720 1725  
 His Arg Asn Arg Leu Leu Pro Gln Ser Lys Pro Arg Leu Leu Pro Pro  
 1730 1735 1740  
 Leu Gly Glu Glu Asp Glu Asp Val Ala Gln Glu Arg Glu Arg Val Thr  
 1745 1750 1755 1760

Lys Gly Ala Thr Gln Gly Asp Val Leu Val Leu Arg Asp Leu Thr Lys  
 1765 1770 1775  
 Val Tyr Arg Gly Gln Arg Asn Pro Ala Val Asp Arg Leu Cys Leu Gly  
 1780 1785 1790  
 Ile Pro Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly  
 1795 1800 1805  
 Lys Thr Ser Thr Phe Arg Met Val Thr Gly Asp Thr Leu Pro Ser Ser  
 1810 1815 1820  
 Gly Glu Ala Val Leu Ala Gly His Asn Val Ala Gln Glu Arg Ser Ala  
 1825 1830 1835 1840  
 Ala His Arg Ser Met Gly Tyr Cys Pro Gln Ser Asp Ala Ile Phe Asp  
 1845 1850 1855  
 Leu Leu Thr Gly Arg Glu His Leu Glu Leu Phe Ala Arg Leu Arg Gly  
 1860 1865 1870  
 Val Pro Glu Ala Gln Val Ala Gln Thr Ala Leu Ser Gly Leu Val Arg  
 1875 1880 1885  
 Leu Gly Leu Pro Ser Tyr Ala Asp Arg Pro Ala Gly Thr Tyr Ser Gly  
 1890 1895 1900  
 Gly Asn Lys Arg Lys Leu Ala Thr Ala Leu Ala Leu Val Gly Asp Pro  
 1905 1910 1915 1920  
 Ala Val Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Ser Ala  
 1925 1930 1935  
 Arg Arg Phe Leu Trp Asn Ser Leu Leu Ser Val Val Arg Glu Gly Arg  
 1940 1945 1950  
 Ser Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys  
 1955 1960 1965  
 Thr Arg Leu Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu Gly Ser  
 1970 1975 1980  
 Ser Gln His Leu Lys Gly Arg Phe Gly Ala Gly His Thr Leu Thr Leu  
 1985 1990 1995 2000  
 Arg Val Pro Pro Asp Gln Pro Glu Pro Ala Ile Ala Phe Ile Arg Ile  
 2005 2010 2015  
 Thr Phe Pro Gly Ala Glu Leu Arg Glu Val His Gly Ser Arg Leu Arg  
 2020 2025 2030  
 Phe Gln Leu Pro Pro Gly Gly Arg Cys Thr Leu Thr Arg Val Phe Arg  
 2035 2040 2045  
 Glu Leu Ala Ala Gln Gly Arg Ala His Gly Val Glu Asp Phe Ser Val  
 2050 2055 2060  
 Ser Gln Thr Thr Leu Glu Glu Val Phe Leu Tyr Phe Ser Lys Asp Gln

25/93

| 90                                                                                                                                                    | 95 | 100 |      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------|----|-----|------|
| ctg cta gcc gat gcc cgc act gtg ctg gga ggg gcc agt gcc cac agg<br>Leu Leu Ala Asp Ala Arg Thr Val Leu Gly Gly Ala Ser Ala His Arg<br>105 110 115 120 |    |     | 569  |
| acg ctg gct ggc cta ggg aag ctg atc gcc acg ctg agg gct gca cgc<br>Thr Leu Ala Gly Leu Gly Lys Leu Ile Ala Thr Leu Arg Ala Ala Arg<br>125 130 135     |    |     | 617  |
| agc acg gcc cag cct caa cca acc aag cag tct cca ctg gaa cca ccc<br>Ser Thr Ala Gln Pro Gln Pro Thr Lys Gln Ser Pro Leu Glu Pro Pro<br>140 145 150     |    |     | 665  |
| atg ctg gat gtc gcg gag ctg ctg acg tca ctg ctg cgc acg gaa tcc<br>Met Leu Asp Val Ala Glu Leu Leu Thr Ser Leu Leu Arg Thr Glu Ser<br>155 160 165     |    |     | 713  |
| ctg ggg ttg gca ctg ggc caa gcc cag gag ccc ttg cac agc ttg ttg<br>Leu Gly Leu Ala Leu Gly Gln Ala Gln Glu Pro Leu His Ser Leu Leu<br>170 175 180     |    |     | 761  |
| gag gcc gct gag gac ctg gcc cag gag ctc ctg gcg ctg cgc agc ctg<br>Glu Ala Ala Glu Asp Leu Ala Gln Glu Leu Leu Ala Leu Arg Ser Leu<br>185 190 195 200 |    |     | 809  |
| gtg gag ctt cgg gca ctg ctg cag aga ccc cga ggg acc agc ggc ccc<br>Val Glu Leu Arg Ala Leu Leu Gln Arg Pro Arg Gly Thr Ser Gly Pro<br>205 210 215     |    |     | 857  |
| ctg gag ttg ctg tca gag gcc ctc tgc agt gtc agg gga cct agc agc<br>Leu Glu Leu Leu Ser Glu Ala Leu Cys Ser Val Arg Gly Pro Ser Ser<br>220 225 230     |    |     | 905  |
| aca gtg ggc ccc tcc ctc aac tgg tac gag gct agt gac ctg atg gag<br>Thr Val Gly Pro Ser Leu Asn Trp Tyr Glu Ala Ser Asp Leu Met Glu<br>235 240 245     |    |     | 953  |
| ctg gtg ggg cag gag cca gaa tcc gcc ctg cca gac agc agc ctg agc<br>Leu Val Gly Gln Glu Pro Glu Ser Ala Leu Pro Asp Ser Ser Leu Ser<br>250 255 260     |    |     | 1001 |
| ccc gcc tgc tgc gag ctg att gga gcc ctg gac agc cac ccg ctg tcc<br>Pro Ala Cys Ser Glu Leu Ile Gly Ala Leu Asp Ser His Pro Leu Ser<br>265 270 275 280 |    |     | 1049 |
| cgc ctg ctc tgg aga cgc ctg aag cct ctg atc ctc ggg aag cta ctc<br>Arg Leu Leu Trp Arg Arg Leu Lys Pro Leu Ile Leu Gly Lys Leu Leu<br>285 290 295     |    |     | 1097 |
| ttt gca cca gat aca cct ttt acc cgg aag ctc atg gcc cag gtg aac<br>Phe Ala Pro Asp Thr Pro Phe Thr Arg Lys Leu Met Ala Gln Val Asn<br>300 305 310     |    |     | 1145 |
| cgg acc ttc gag gag ctc acc ctg ctg agg gat gtc cgg gag gtg tgg<br>Arg Thr Phe Glu Glu Leu Thr Leu Leu Arg Asp Val Arg Glu Val Trp<br>315 320 325     |    |     | 1193 |

|                                                                 |      |
|-----------------------------------------------------------------|------|
| gag atg ctg gga ccc cgg atc ttc acc ttc atg aac gac agt tcc aat | 1241 |
| Glu Met Leu Gly Pro Arg Ile Phe Thr Phe Met Asn Asp Ser Ser Asn |      |
| 330 335 340                                                     |      |
| gtg gcc atg ctg cag cgg ctc ctg cag atg cag gat gaa gga aga agg | 1289 |
| Val Ala Met Leu Gln Arg Leu Leu Gln Met Gln Asp Glu Gly Arg Arg |      |
| 345 350 355 360                                                 |      |
| cag ccc aga cct gga ggc cgg gac cac atg gag gcc ctg cga tcc ttt | 1337 |
| Gln Pro Arg Pro Gly Gly Arg Asp His Met Glu Ala Leu Arg Ser Phe |      |
| 365 370 375                                                     |      |
| ctg gac cct ggg agc ggt ggc tac agc tgg cag gac gca cac gct gat | 1385 |
| Leu Asp Pro Gly Ser Gly Gly Tyr Ser Trp Gln Asp Ala His Ala Asp |      |
| 380 385 390                                                     |      |
| gtg ggg cac ctg gtg ggc acg ctg ggc cga gtg acg gag tgc ctg tcc | 1433 |
| Val Gly His Leu Val Gly Thr Leu Gly Arg Val Thr Glu Cys Leu Ser |      |
| 395 400 405                                                     |      |
| ttg gac aag ctg gag gcg gca ccc tca gag gca gcc ctg gtg tcg cgg | 1481 |
| Leu Asp Lys Leu Glu Ala Ala Pro Ser Glu Ala Ala Leu Val Ser Arg |      |
| 410 415 420                                                     |      |
| gcc ctg caa ctg ctc gcg gaa cat cga ttc tgg gcc ggc gtc gtc ttc | 1529 |
| Ala Leu Gln Leu Leu Ala Glu His Arg Phe Trp Ala Gly Val Val Phe |      |
| 425 430 435 440                                                 |      |
| ttg gga cct gag gac tct tca gac ccc aca gag cac cca acc cca gac | 1577 |
| Leu Gly Pro Glu Asp Ser Ser Asp Pro Thr Glu His Pro Thr Pro Asp |      |
| 445 450 455                                                     |      |
| ctg ggc ccc ggc cac gtg cgc atc aaa atc cgc atg gac att gac gtg | 1625 |
| Leu Gly Pro Gly His Val Arg Ile Lys Ile Arg Met Asp Ile Asp Val |      |
| 460 465 470                                                     |      |
| gtc acg agg acc aat aag atc agg gac agg ttt tgg gac cct ggc cca | 1673 |
| Val Thr Arg Thr Asn Lys Ile Arg Asp Arg Phe Trp Asp Pro Gly Pro |      |
| 475 480 485                                                     |      |
| gcc gcg gac ccc ctg acc gac ctg cgc tac gtg tgg ggc ggc ttc gtg | 1721 |
| Ala Ala Asp Pro Leu Thr Asp Leu Arg Tyr Val Trp Gly Gly Phe Val |      |
| 490 495 500                                                     |      |
| tac ctg caa gac ctg gtg gag cgt gca gcc gtc cgc gtg ctc agc ggc | 1769 |
| Tyr Leu Gln Asp Leu Val Glu Arg Ala Ala Val Arg Val Leu Ser Gly |      |
| 505 510 515 520                                                 |      |
| gcc aac ccc cgg gcc ggc ctc tac ctg cag cag atg ccc tat ccg tgc | 1817 |
| Ala Asn Pro Arg Ala Gly Leu Tyr Leu Gln Gln Met Pro Tyr Pro Cys |      |
| 525 530 535                                                     |      |
| tat gtg gac gac gtg ttc ctg cgt gtg ctg agc cgg tcg ctg ccg ctc | 1865 |
| Tyr Val Asp Asp Val Phe Leu Arg Val Leu Ser Arg Ser Leu Pro Leu |      |
| 540 545 550                                                     |      |
| ttc ctg acg ctg gcc tgg atc tac tcc gtg aca ctg aca gtg aag gcc | 1913 |
| Phe Leu Thr Leu Ala Trp Ile Tyr Ser Val Thr Leu Thr Val Lys Ala |      |

28/93

|     |     |     |     |      |     |     |     |     |      |     |     |     |     |      |      |      |
|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|-----|------|------|------|
| aag | gtg | ctg | gta | gaa  | gag | gca | ccg | ccc | ggc  | ctg | agt | cct | ggc | gta  | tcc  | 2633 |
| Lys | Val | Leu | Val | Glu  | Glu | Ala | Pro | Pro | Gly  | Leu | Ser | Pro | Gly | Val  | Ser  |      |
|     |     | 795 |     |      |     |     | 800 |     |      |     |     | 805 |     |      |      |      |
| gtt | cgc | agc | ctg | gag  | aag | cgc | ttt | cct | gga  | agc | ccg | cag | cca | gcc  | ctg  | 2681 |
| Val | Arg | Ser | Leu | Glu  | Lys | Arg | Phe | Pro | Gly  | Ser | Pro | Gln | Pro | Ala  | Leu  |      |
|     | 810 |     |     |      |     | 815 |     |     |      |     | 820 |     |     |      |      |      |
| cgg | ggg | ctc | agc | ctg  | gac | ttc | tac | cag | ggc  | cac | atc | acc | gcc | ttc  | ctg  | 2729 |
| Arg | Gly | Leu | Ser | Leu  | Asp | Phe | Tyr | Gln | Gly  | His | Ile | Thr | Ala | Phe  | Leu  |      |
| 825 |     |     |     |      | 830 |     |     |     |      | 835 |     |     |     |      | 840  |      |
| ggc | cac | aac | ggg | gcc  | ggc | aag | acc | acc | acc  | ctg | tcc | atc | ttg | agt  | ggc  | 2777 |
| Gly | His | Asn | Gly | Ala  | Gly | Lys | Thr | Thr | Thr  | Leu | Ser | Ile | Leu | Ser  | Gly  |      |
|     |     |     | 845 |      |     |     |     |     | 850  |     |     |     |     | 855  |      |      |
| ctc | ttc | cca | ccc | agt  | ggt | ggc | tct | gcc | ttc  | atc | ctg | ggc | cac | gac  | gtc  | 2825 |
| Leu | Phe | Pro | Pro | Ser  | Gly | Gly | Ser | Ala | Phe  | Ile | Leu | Gly | His | Asp  | Val  |      |
|     |     | 860 |     |      |     |     |     | 865 |      |     |     |     | 870 |      |      |      |
| cgc | tcc | agc | atg | gcc  | gcc | atc | cgg | ccc | cac  | ctg | ggc | gtc | tgt | cct  | cag  | 2873 |
| Arg | Ser | Ser | Met | Ala  | Ala | Ile | Arg | Pro | His  | Leu | Gly | Val | Cys | Pro  | Gln  |      |
|     |     | 875 |     |      |     |     | 880 |     |      |     |     | 885 |     |      |      |      |
| tac | aac | gtg | ctg | ttt  | gac | atg | ctg | acc | gtg  | gac | gag | cac | gtc | tgg  | ttc  | 2921 |
| Tyr | Asn | Val | Leu | Phe  | Asp | Met | Leu | Thr | Val  | Asp | Glu | His | Val | Trp  | Phe  |      |
|     | 890 |     |     |      |     | 895 |     |     |      |     | 900 |     |     |      |      |      |
| tat | ggg | cgg | ctg | aag  | ggt | ctg | agt | gcc | gct  | gta | gtg | ggc | ccc | gag  | cag  | 2969 |
| Tyr | Gly | Arg | Leu | Lys  | Gly | Leu | Ser | Ala | Ala  | Val | Val | Gly | Pro | Glu  | Gln  |      |
| 905 |     |     |     |      | 910 |     |     |     |      | 915 |     |     |     |      | 920  |      |
| gac | cgt | ctg | ctg | cag  | gat | gtg | ggg | ctg | gtc  | tcc | aag | cag | agt | gtg  | cag  | 3017 |
| Asp | Arg | Leu | Leu | Gln  | Asp | Val | Gly | Leu | Val  | Ser | Lys | Gln | Ser | Val  | Gln  |      |
|     |     |     |     | 925  |     |     |     |     | 930  |     |     |     |     | 935  |      |      |
| act | cgc | cac | ctc | tct  | ggt | ggg | atg | caa | cgg  | aag | ctg | tcc | gtg | gcc  | att  | 3065 |
| Thr | Arg | His | Leu | Ser  | Gly | Gly | Met | Gln | Arg  | Lys | Leu | Ser | Val | Ala  | Ile  |      |
|     |     |     | 940 |      |     |     |     | 945 |      |     |     |     | 950 |      |      |      |
| gcc | ttt | gtg | ggc | ggc  | tcc | caa | gtt | gtt | atc  | ctg | gac | gag | cct | acg  | gct  | 3113 |
| Ala | Phe | Val | Gly | Gly  | Ser | Gln | Val | Val | Ile  | Leu | Asp | Glu | Pro | Thr  | Ala  |      |
|     |     | 955 |     |      |     |     | 960 |     |      |     |     | 965 |     |      |      |      |
| ggc | gtg | gat | cct | gct  | tcc | cgc | cgc | ggt | att  | tgg | gag | ctg | ctg | ctc  | aaa  | 3161 |
| Gly | Val | Asp | Pro | Ala  | Ser | Arg | Arg | Gly | Ile  | Trp | Glu | Leu | Leu | Leu  | Lys  |      |
|     |     | 970 |     |      |     | 975 |     |     |      |     | 980 |     |     |      |      |      |
| tac | cga | gaa | ggt | cgc  | acg | ctg | atc | ctc | tcc  | acc | cac | cac | ctg | gat  | gag  | 3209 |
| Tyr | Arg | Glu | Gly | Arg  | Thr | Leu | Ile | Leu | Ser  | Thr | His | His | Leu | Asp  | Glu  |      |
| 985 |     |     |     |      | 990 |     |     |     |      | 995 |     |     |     |      | 1000 |      |
| gca | gag | ctg | ctg | gga  | gac | cgt | gtg | gct | gtg  | gtg | gca | ggt | ggc | cgc  | ttg  | 3257 |
| Ala | Glu | Leu | Leu | Gly  | Asp | Arg | Val | Ala | Val  | Val | Ala | Gly | Gly | Arg  | Leu  |      |
|     |     |     |     | 1005 |     |     |     |     | 1010 |     |     |     |     | 1015 |      |      |
| tgc | tgc | tgt | ggc | tcc  | cca | ctc | ttc | ctg | cgc  | cgt | cac | ctg | ggc | tcc  | ggc  | 3305 |
| Cys | Cys | Cys | Gly | Ser  | Pro | Leu | Phe | Leu | Arg  | Arg | His | Leu | Gly | Ser  | Gly  |      |

| 1020                                                            | 1025 | 1030 |      |
|-----------------------------------------------------------------|------|------|------|
| tac tac ctg acg ctg gtg aag gcc cgc ctg ccc ctg acc acc aat gag |      |      | 3353 |
| Tyr Tyr Leu Thr Leu Val Lys Ala Arg Leu Pro Leu Thr Thr Asn Glu |      |      |      |
| 1035                                                            | 1040 | 1045 |      |
| aag gct gac act gac atg gag ggc agt gtg gac acc agg cag gaa aag |      |      | 3401 |
| Lys Ala Asp Thr Asp Met Glu Gly Ser Val Asp Thr Arg Gln Glu Lys |      |      |      |
| 1050                                                            | 1055 | 1060 |      |
| aag aat ggc agc cag ggc agc aga gtc ggc act cct cag ctg ctg gcc |      |      | 3449 |
| Lys Asn Gly Ser Gln Gly Ser Arg Val Gly Thr Pro Gln Leu Leu Ala |      |      |      |
| 1065                                                            | 1070 | 1075 | 1080 |
| ctg gta cag cac tgg gtg ccc ggg gca cgg ctg gtg gag gag ctg cca |      |      | 3497 |
| Leu Val Gln His Trp Val Pro Gly Ala Arg Leu Val Glu Glu Leu Pro |      |      |      |
| 1085                                                            | 1090 | 1095 |      |
| cac gag ctg gtg ctg gtg ctg ccc tac acg ggt gcc cat gac ggc agc |      |      | 3545 |
| His Glu Leu Val Leu Val Leu Pro Tyr Thr Gly Ala His Asp Gly Ser |      |      |      |
| 1100                                                            | 1105 | 1110 |      |
| ttc gcc aca ctc ttc cga gag cta gac acg cgg ctg gcg gag ctg agg |      |      | 3593 |
| Phe Ala Thr Leu Phe Arg Glu Leu Asp Thr Arg Leu Ala Glu Leu Arg |      |      |      |
| 1115                                                            | 1120 | 1125 |      |
| ctc act ggc tac ggg atc tcc gac acc agc ctc gag gag atc ttc ctg |      |      | 3641 |
| Leu Thr Gly Tyr Gly Ile Ser Asp Thr Ser Leu Glu Glu Ile Phe Leu |      |      |      |
| 1130                                                            | 1135 | 1140 |      |
| aag gtg gtg gag gag tgt gct gcg gac aca gat atg gag gat ggc agc |      |      | 3689 |
| Lys Val Val Glu Glu Cys Ala Ala Asp Thr Asp Met Glu Asp Gly Ser |      |      |      |
| 1145                                                            | 1150 | 1155 | 1160 |
| tgc ggg cag cac cta tgc aca ggc att gct ggc cta gac gta acc ctg |      |      | 3737 |
| Cys Gly Gln His Leu Cys Thr Gly Ile Ala Gly Leu Asp Val Thr Leu |      |      |      |
| 1165                                                            | 1170 | 1175 |      |
| cgg ctc aag atg ccg cca cag gag aca gcg ctg gag aac ggg gaa cca |      |      | 3785 |
| Arg Leu Lys Met Pro Pro Gln Glu Thr Ala Leu Glu Asn Gly Glu Pro |      |      |      |
| 1180                                                            | 1185 | 1190 |      |
| gct ggg tca gcc cca gag act gac cag ggc tct ggg cca gac gcc gtg |      |      | 3833 |
| Ala Gly Ser Ala Pro Glu Thr Asp Gln Gly Ser Gly Pro Asp Ala Val |      |      |      |
| 1195                                                            | 1200 | 1205 |      |
| ggc cgg gta cag ggc tgg gca ctg acc cgc cag cag ctc cag gcc ctg |      |      | 3881 |
| Gly Arg Val Gln Gly Trp Ala Leu Thr Arg Gln Gln Leu Gln Ala Leu |      |      |      |
| 1210                                                            | 1215 | 1220 |      |
| ctt ctc aag cgc ttt ctg ctt gcc cgc cgc agc cgc cgc ggc ctg ttc |      |      | 3929 |
| Leu Leu Lys Arg Phe Leu Leu Ala Arg Arg Ser Arg Arg Gly Leu Phe |      |      |      |
| 1225                                                            | 1230 | 1235 | 1240 |
| gcc cag atc gtg ctg cct gcc ctc ttt gtg ggc ctg gcc ctc gtg ttc |      |      | 3977 |
| Ala Gln Ile Val Leu Pro Ala Leu Phe Val Gly Leu Ala Leu Val Phe |      |      |      |
| 1245                                                            | 1250 | 1255 |      |

|                                                                 |      |
|-----------------------------------------------------------------|------|
| agc ctc atc gtg cct cct ttc ggg cac tac ccg gct ctg cgg ctc agt | 4025 |
| Ser Leu Ile Val Pro Pro Phe Gly His Tyr Pro Ala Leu Arg Leu Ser |      |
| 1260 1265 1270                                                  |      |
| ccc acc atg tac ggt gct cag gtg tcc ttc ttc agt gag gac gcc cca | 4073 |
| Pro Thr Met Tyr Gly Ala Gln Val Ser Phe Phe Ser Glu Asp Ala Pro |      |
| 1275 1280 1285                                                  |      |
| ggg gac cct gga cgt gcc cgg ctg ctc gag gcg ctg ctg cag gag gca | 4121 |
| Gly Asp Pro Gly Arg Ala Arg Leu Leu Glu Ala Leu Leu Gln Glu Ala |      |
| 1290 1295 1300                                                  |      |
| gga ctg gag gag ccc cca gtg cag cat agc tcc cac agg ttc tcg gca | 4169 |
| Gly Leu Glu Glu Pro Pro Val Gln His Ser Ser His Arg Phe Ser Ala |      |
| 1305 1310 1315 1320                                             |      |
| cca gaa gtt cct gct gaa gtg gcc aag gtc ttg gcc agt ggc aac tgg | 4217 |
| Pro Glu Val Pro Ala Glu Val Ala Lys Val Leu Ala Ser Gly Asn Trp |      |
| 1325 1330 1335                                                  |      |
| acc cca gag tct cca tcc cca gcc tgc cag tgt agc cag ccc ggt gcc | 4265 |
| Thr Pro Glu Ser Pro Ser Pro Ala Cys Gln Cys Ser Gln Pro Gly Ala |      |
| 1340 1345 1350                                                  |      |
| cgg cgc ctg ctg ccc gac tgc ccg gct gca gct ggt ggt ccc cct ccg | 4313 |
| Arg Arg Leu Leu Pro Asp Cys Pro Ala Ala Ala Gly Gly Pro Pro Pro |      |
| 1355 1360 1365                                                  |      |
| ccc cag gca gtg acc ggc tct ggg gaa gtg gtt cag aac ctg aca ggc | 4361 |
| Pro Gln Ala Val Thr Gly Ser Gly Glu Val Val Gln Asn Leu Thr Gly |      |
| 1370 1375 1380                                                  |      |
| cgg aac ctg tct gac ttc ctg gtc aag acc tac ccg cgc ctg gtg cgc | 4409 |
| Arg Asn Leu Ser Asp Phe Leu Val Lys Thr Tyr Pro Arg Leu Val Arg |      |
| 1385 1390 1395 1400                                             |      |
| cag ggc ctg aag act aag aag tgg gtg aat gag gtc agg tac gga ggc | 4457 |
| Gln Gly Leu Lys Thr Lys Lys Trp Val Asn Glu Val Arg Tyr Gly Gly |      |
| 1405 1410 1415                                                  |      |
| ttc tcg ctg ggg ggc cga gac cca ggc ctg ccc tcg ggc caa gag ttg | 4505 |
| Phe Ser Leu Gly Gly Arg Asp Pro Gly Leu Pro Ser Gly Gln Glu Leu |      |
| 1420 1425 1430                                                  |      |
| ggc cgc tca gtg gag gag ttg tgg gcg ctg ctg agt ccc ctg cct ggc | 4553 |
| Gly Arg Ser Val Glu Glu Leu Trp Ala Leu Leu Ser Pro Leu Pro Gly |      |
| 1435 1440 1445                                                  |      |
| ggg gcc ctc gac cgt gtc ctg aaa aac ctc aca gcc tgg gct cac agc | 4601 |
| Gly Ala Leu Asp Arg Val Leu Lys Asn Leu Thr Ala Trp Ala His Ser |      |
| 1450 1455 1460                                                  |      |
| ctg gat gct cag gac agt ctc aag atc tgg ttc aac aac aaa ggc tgg | 4649 |
| Leu Asp Ala Gln Asp Ser Leu Lys Ile Trp Phe Asn Asn Lys Gly Trp |      |
| 1465 1470 1475 1480                                             |      |
| cac tcc atg gtg gcc ttt gtc aac cga gcc agc aac gca atc ctc cgt | 4697 |
| His Ser Met Val Ala Phe Val Asn Arg Ala Ser Asn Ala Ile Leu Arg |      |

|  | 1485                                                            | 1490 | 1495 |      |
|--|-----------------------------------------------------------------|------|------|------|
|  | gct cac ctg ccc cca ggc ccg gcc cgc cac gcc cac agc atc acc aca |      |      | 4745 |
|  | Ala His Leu Pro Pro Gly Pro Ala Arg His Ala His Ser Ile Thr Thr |      |      |      |
|  | 1500                                                            | 1505 | 1510 |      |
|  | ctc aac cac ccc ttg aac ctc acc aag gag cag ctg tct gag gct gca |      |      | 4793 |
|  | Leu Asn His Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ala Ala |      |      |      |
|  | 1515                                                            | 1520 | 1525 |      |
|  | ctg atg gcc tcc tcg gtg gac gtc ctc gtc tcc atc tgt gtg gtc ttt |      |      | 4841 |
|  | Leu Met Ala Ser Ser Val Asp Val Leu Val Ser Ile Cys Val Val Phe |      |      |      |
|  | 1530                                                            | 1535 | 1540 |      |
|  | gcc atg tcc ttt gtc ccg gcc agc ttc act ctt gtc ctc att gag gag |      |      | 4889 |
|  | Ala Met Ser Phe Val Pro Ala Ser Phe Thr Leu Val Leu Ile Glu Glu |      |      |      |
|  | 1545                                                            | 1550 | 1555 | 1560 |
|  | cga gtc acc cga gcc aag cac ctg cag ctc atg ggg ggc ctg tcc ccc |      |      | 4937 |
|  | Arg Val Thr Arg Ala Lys His Leu Gln Leu Met Gly Gly Leu Ser Pro |      |      |      |
|  | 1565                                                            | 1570 | 1575 |      |
|  | acc ctc tac tgg ctt ggc aac ttt ctc tgg gac atg tgt aac tac ttg |      |      | 4985 |
|  | Thr Leu Tyr Trp Leu Gly Asn Phe Leu Trp Asp Met Cys Asn Tyr Leu |      |      |      |
|  | 1580                                                            | 1585 | 1590 |      |
|  | gtg cca gca tgc atc gtg gtg ctc atc ttt ctg gcc ttc cag cag agg |      |      | 5033 |
|  | Val Pro Ala Cys Ile Val Val Leu Ile Phe Leu Ala Phe Gln Gln Arg |      |      |      |
|  | 1595                                                            | 1600 | 1605 |      |
|  | gca tat gtg gcc cct gcc aac ctg cct gct ctc ctg ctg ttg cta cta |      |      | 5081 |
|  | Ala Tyr Val Ala Pro Ala Asn Leu Pro Ala Leu Leu Leu Leu Leu     |      |      |      |
|  | 1610                                                            | 1615 | 1620 |      |
|  | ctg tat ggc tgg tcg atc aca ccg ctc atg tac cca gcc tcc ttc ttc |      |      | 5129 |
|  | Leu Tyr Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala Ser Phe Phe |      |      |      |
|  | 1625                                                            | 1630 | 1635 | 1640 |
|  | ttc tcc gtg ccc agc aca gcc tat gtg gtg ctc acc tgc ata aac ctc |      |      | 5177 |
|  | Phe Ser Val Pro Ser Thr Ala Tyr Val Val Leu Thr Cys Ile Asn Leu |      |      |      |
|  | 1645                                                            | 1650 | 1655 |      |
|  | ttt att ggc atc aat gga agc atg gcc acc ttt gtg ctt gag ctc ttc |      |      | 5225 |
|  | Phe Ile Gly Ile Asn Gly Ser Met Ala Thr Phe Val Leu Glu Leu Phe |      |      |      |
|  | 1660                                                            | 1665 | 1670 |      |
|  | tct gat cag aag ctg cag gag gtg agc cgg atc ttg aaa cag gtc ttc |      |      | 5273 |
|  | Ser Asp Gln Lys Leu Gln Glu Val Ser Arg Ile Leu Lys Gln Val Phe |      |      |      |
|  | 1675                                                            | 1680 | 1685 |      |
|  | ctt atc ttc ccc cac ttc tgc ttg ggc cgg ggg ctc att gac atg gtg |      |      | 5321 |
|  | Leu Ile Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Met Val |      |      |      |
|  | 1690                                                            | 1695 | 1700 |      |
|  | cgg aac cag gcc atg gct gat gcc ttt gag cgc ttg gga gac agg cag |      |      | 5369 |
|  | Arg Asn Gln Ala Met Ala Asp Ala Phe Glu Arg Leu Gly Asp Arg Gln |      |      |      |
|  | 1705                                                            | 1710 | 1715 | 1720 |

|                                                                 |      |
|-----------------------------------------------------------------|------|
| ttc cag tca ccc ctg cgc tgg gag gtg gtc ggc aag aac ctc ttg gcc | 5417 |
| Phe Gln Ser Pro Leu Arg Trp Glu Val Val Gly Lys Asn Leu Leu Ala |      |
| 1725 1730 1735                                                  |      |
| atg gtg ata cag ggg ccc ctc ttc ctt ctc ttc aca cta ctg ctg cag | 5465 |
| Met Val Ile Gln Gly Pro Leu Phe Leu Leu Phe Thr Leu Leu Leu Gln |      |
| 1740 1745 1750                                                  |      |
| cac cga agc caa ctc ctg cca cag ccc agg gtg agg tct ctg cca ctc | 5513 |
| His Arg Ser Gln Leu Leu Pro Gln Pro Arg Val Arg Ser Leu Pro Leu |      |
| 1755 1760 1765                                                  |      |
| ctg gga gag gag gac gag gat gta gcc cgt gaa cgg gag cgg gtg gtc | 5561 |
| Leu Gly Glu Glu Asp Glu Asp Val Ala Arg Glu Arg Glu Arg Val Val |      |
| 1770 1775 1780                                                  |      |
| caa gga gcc acc cag ggg gat gtg ttg gtg ctg agg aac ttg acc aag | 5609 |
| Gln Gly Ala Thr Gln Gly Asp Val Leu Val Leu Arg Asn Leu Thr Lys |      |
| 1785 1790 1795 1800                                             |      |
| gta tac cgt ggg cag agg atg cca gct gtt gac cgc ttg tgc ctg ggg | 5657 |
| Val Tyr Arg Gly Gln Arg Met Pro Ala Val Asp Arg Leu Cys Leu Gly |      |
| 1805 1810 1815                                                  |      |
| att ccc cct ggt gag tgt ttt ggg ctg ctg ggt gtg aat gga gca ggg | 5705 |
| Ile Pro Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly |      |
| 1820 1825 1830                                                  |      |
| aag acg tcc acg ttt cgc atg gtg acg ggg gac aca ttg gcc agc agg | 5753 |
| Lys Thr Ser Thr Phe Arg Met Val Thr Gly Asp Thr Leu Ala Ser Arg |      |
| 1835 1840 1845                                                  |      |
| ggc gag gct gtg ctg gca ggc cac agc gtg gcc cgg gaa ccc agt gct | 5801 |
| Gly Glu Ala Val Leu Ala Gly His Ser Val Ala Arg Glu Pro Ser Ala |      |
| 1850 1855 1860                                                  |      |
| gcg cac ctc agc atg gga tac tgc cct caa tcc gat gcc atc ttt gag | 5849 |
| Ala His Leu Ser Met Gly Tyr Cys Pro Gln Ser Asp Ala Ile Phe Glu |      |
| 1865 1870 1875 1880                                             |      |
| ctg ctg acg ggc cgc gag cac ctg gag ctg ctt gcg cgc ctg cgc ggt | 5897 |
| Leu Leu Thr Gly Arg Glu His Leu Glu Leu Leu Ala Arg Leu Arg Gly |      |
| 1885 1890 1895                                                  |      |
| gtc ccg gag gcc cag gtt gcc cag acc gct ggc tca ggc ctg gcg cgt | 5945 |
| Val Pro Glu Ala Gln Val Ala Gln Thr Ala Gly Ser Gly Leu Ala Arg |      |
| 1900 1905 1910                                                  |      |
| ctg gga ctc tca tgg tac gca gac cgg cct gca ggc acc tac agc gga | 5993 |
| Leu Gly Leu Ser Trp Tyr Ala Asp Arg Pro Ala Gly Thr Tyr Ser Gly |      |
| 1915 1920 1925                                                  |      |
| ggg aac aaa cgc aag ctg gcg acg gcc ctg gcg ctg gtt ggg gac cca | 6041 |
| Gly Asn Lys Arg Lys Leu Ala Thr Ala Leu Ala Leu Val Gly Asp Pro |      |
| 1930 1935 1940                                                  |      |
| gcc gtg gtg ttt ctg gac gag ccg acc aca ggc atg gac ccc agc gcg | 6089 |
| Ala Val Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Ser Ala |      |

| 1945                                                              | 1950 | 1955 | 1960 |      |
|-------------------------------------------------------------------|------|------|------|------|
| cgg cgc ttc ctt tgg aac agc ctt ttg gcc gtg gtg cgg gag ggc cgt   |      |      |      | 6137 |
| Arg Arg Phe Leu Trp Asn Ser Leu Leu Ala Val Val Arg Glu Gly Arg   |      |      |      |      |
|                                                                   | 1965 | 1970 | 1975 |      |
| tca gtg atg ctc acc tcc cat agc atg gag gag tgt gaa gcg ctc tgc   |      |      |      | 6185 |
| Ser Val Met Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys   |      |      |      |      |
|                                                                   | 1980 | 1985 | 1990 |      |
| tcg cgc cta gcc atc atg gtg aat ggg cgg ttc cgc tgc ctg ggc agc   |      |      |      | 6233 |
| Ser Arg Leu Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu Gly Ser   |      |      |      |      |
|                                                                   | 1995 | 2000 | 2005 |      |
| ccg caa cat ctc aag ggc aga ttc gcg gcg ggt cac aca ctg acc ctg   |      |      |      | 6281 |
| Pro Gln His Leu Lys Gly Arg Phe Ala Ala Gly His Thr Leu Thr Leu   |      |      |      |      |
|                                                                   | 2010 | 2015 | 2020 |      |
| cgg gtg ccc gcc gca agg tcc cag ccg gca gcg gcc ttc gtg gcg gcc   |      |      |      | 6329 |
| Arg Val Pro Ala Ala Arg Ser Gln Pro Ala Ala Ala Phe Val Ala Ala   |      |      |      |      |
|                                                                   | 2025 | 2030 | 2035 | 2040 |
| gag ttc cct ggg tcg gag ctg cgc gag gca cat gga ggc cgc ctg cgc   |      |      |      | 6377 |
| Glu Phe Pro Gly Ser Glu Leu Arg Glu Ala His Gly Gly Arg Leu Arg   |      |      |      |      |
|                                                                   | 2045 | 2050 | 2055 |      |
| ttc cag ctg ccg ccg gga ggg cgc tgc gcc ctg gcg cgc gtc ttt gga   |      |      |      | 6425 |
| Phe Gln Leu Pro Pro Gly Gly Arg Cys Ala Leu Ala Arg Val Phe Gly   |      |      |      |      |
|                                                                   | 2060 | 2065 | 2070 |      |
| gag ctg gcg gtg cac ggc gca gag cac ggc gtg gag gac ttt tcc gtg   |      |      |      | 6473 |
| Glu Leu Ala Val His Gly Ala Glu His Gly Val Glu Asp Phe Ser Val   |      |      |      |      |
|                                                                   | 2075 | 2080 | 2085 |      |
| agc cag acg atg ctg gag gag gta ttc ttg tac ttc tcc aag gac cag   |      |      |      | 6521 |
| Ser Gln Thr Met Leu Glu Glu Val Phe Leu Tyr Phe Ser Lys Asp Gln   |      |      |      |      |
|                                                                   | 2090 | 2095 | 2100 |      |
| ggg aag gac gag gac acc gaa gag cag aag gag gca gga gtg gga gtg   |      |      |      | 6569 |
| Gly Lys Asp Glu Asp Thr Glu Glu Gln Lys Glu Ala Gly Val Gly Val   |      |      |      |      |
|                                                                   | 2105 | 2110 | 2115 | 2120 |
| gac ccc gcg cca ggc ctg cag cac ccc aaa cgc gtc agc cag ttc ctc   |      |      |      | 6617 |
| Asp Pro Ala Pro Gly Leu Gln His Pro Lys Arg Val Ser Gln Phe Leu   |      |      |      |      |
|                                                                   | 2125 | 2130 | 2135 |      |
| gat gac cct agc act gcc gag act gtg ctc tga gcctccctcc cctgcggggc |      |      |      | 6670 |
| Asp Asp Pro Ser Thr Ala Glu Thr Val Leu                           |      |      |      |      |
|                                                                   | 2140 | 2145 |      |      |
| cgcggggagg ccctgggaat ggcaagggca aggtagagtg cctaggagcc ctggactcag |      |      |      | 6730 |
| gctggcagag gggctggtgc cctggagaaa ataaagagaa ggctggagag aagccgtgct |      |      |      | 6790 |
| ggtgaaaaaa aaaa                                                   |      |      |      | 6804 |

&lt;210&gt; 5

&lt;211&gt; 2146

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 5

```

Met Ala Phe Trp Thr Gln Leu Met Leu Leu Leu Trp Lys Asn Phe Met
  1              5              10              15

Tyr Arg Arg Arg Gln Pro Val Gln Leu Leu Val Glu Leu Leu Trp Pro
      20              25              30

Leu Phe Leu Phe Phe Ile Leu Val Ala Val Arg His Ser His Pro Pro
      35              40              45

Leu Glu His His Glu Cys His Phe Pro Asn Lys Pro Leu Pro Ser Ala
      50              55              60

Gly Thr Val Pro Trp Leu Gln Gly Leu Ile Cys Asn Val Asn Asn Thr
      65              70              75              80

Cys Phe Pro Gln Leu Thr Pro Gly Glu Glu Pro Gly Arg Leu Ser Asn
      85              90              95

Phe Asn Asp Ser Leu Val Ser Arg Leu Leu Ala Asp Ala Arg Thr Val
      100              105              110

Leu Gly Gly Ala Ser Ala His Arg Thr Leu Ala Gly Leu Gly Lys Leu
      115              120              125

Ile Ala Thr Leu Arg Ala Ala Arg Ser Thr Ala Gln Pro Gln Pro Thr
      130              135              140

Lys Gln Ser Pro Leu Glu Pro Pro Met Leu Asp Val Ala Glu Leu Leu
      145              150              155              160

Thr Ser Leu Leu Arg Thr Glu Ser Leu Gly Leu Ala Leu Gly Gln Ala
      165              170              175

Gln Glu Pro Leu His Ser Leu Leu Glu Ala Ala Glu Asp Leu Ala Gln
      180              185              190

Glu Leu Leu Ala Leu Arg Ser Leu Val Glu Leu Arg Ala Leu Leu Gln
      195              200              205

Arg Pro Arg Gly Thr Ser Gly Pro Leu Glu Leu Leu Ser Glu Ala Leu
      210              215              220

Cys Ser Val Arg Gly Pro Ser Ser Thr Val Gly Pro Ser Leu Asn Trp
      225              230              235              240

Tyr Glu Ala Ser Asp Leu Met Glu Leu Val Gly Gln Glu Pro Glu Ser
      245              250              255

Ala Leu Pro Asp Ser Ser Leu Ser Pro Ala Cys Ser Glu Leu Ile Gly
      260              265              270

Ala Leu Asp Ser His Pro Leu Ser Arg Leu Leu Trp Arg Arg Leu Lys
      275              280              285

```

Pro Leu Ile Leu Gly Lys Leu Leu Phe Ala Pro Asp Thr Pro Phe Thr  
 290 295 300  
 Arg Lys Leu Met Ala Gln Val Asn Arg Thr Phe Glu Glu Leu Thr Leu  
 305 310 315 320  
 Leu Arg Asp Val Arg Glu Val Trp Glu Met Leu Gly Pro Arg Ile Phe  
 325 330 335  
 Thr Phe Met Asn Asp Ser Ser Asn Val Ala Met Leu Gln Arg Leu Leu  
 340 345 350  
 Gln Met Gln Asp Glu Gly Arg Arg Gln Pro Arg Pro Gly Gly Arg Asp  
 355 360 365  
 His Met Glu Ala Leu Arg Ser Phe Leu Asp Pro Gly Ser Gly Gly Tyr  
 370 375 380  
 Ser Trp Gln Asp Ala His Ala Asp Val Gly His Leu Val Gly Thr Leu  
 385 390 395 400  
 Gly Arg Val Thr Glu Cys Leu Ser Leu Asp Lys Leu Glu Ala Ala Pro  
 405 410 415  
 Ser Glu Ala Ala Leu Val Ser Arg Ala Leu Gln Leu Leu Ala Glu His  
 420 425 430  
 Arg Phe Trp Ala Gly Val Val Phe Leu Gly Pro Glu Asp Ser Ser Asp  
 435 440 445  
 Pro Thr Glu His Pro Thr Pro Asp Leu Gly Pro Gly His Val Arg Ile  
 450 455 460  
 Lys Ile Arg Met Asp Ile Asp Val Val Thr Arg Thr Asn Lys Ile Arg  
 465 470 475 480  
 Asp Arg Phe Trp Asp Pro Gly Pro Ala Ala Asp Pro Leu Thr Asp Leu  
 485 490 495  
 Arg Tyr Val Trp Gly Gly Phe Val Tyr Leu Gln Asp Leu Val Glu Arg  
 500 505 510  
 Ala Ala Val Arg Val Leu Ser Gly Ala Asn Pro Arg Ala Gly Leu Tyr  
 515 520 525  
 Leu Gln Gln Met Pro Tyr Pro Cys Tyr Val Asp Asp Val Phe Leu Arg  
 530 535 540  
 Val Leu Ser Arg Ser Leu Pro Leu Phe Leu Thr Leu Ala Trp Ile Tyr  
 545 550 555 560  
 Ser Val Thr Leu Thr Val Lys Ala Val Val Arg Glu Lys Glu Thr Arg  
 565 570 575  
 Leu Arg Asp Thr Met Arg Ala Met Gly Leu Ser Arg Ala Val Leu Trp  
 580 585 590

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

| 900  |      |      |      |      |      |      |      |      |      | 905 |      |      |      |      | 910  |  |  |  |  |
|------|------|------|------|------|------|------|------|------|------|-----|------|------|------|------|------|--|--|--|--|
| Ala  | Ala  | Val  | Val  | Gly  | Pro  | Glu  | Gln  | Asp  | Arg  | Leu | Leu  | Gln  | Asp  | Val  | Gly  |  |  |  |  |
|      |      | 915  |      |      |      |      | 920  |      |      |     |      | 925  |      |      |      |  |  |  |  |
| Leu  | Val  | Ser  | Lys  | Gln  | Ser  | Val  | Gln  | Thr  | Arg  | His | Leu  | Ser  | Gly  | Gly  | Met  |  |  |  |  |
|      |      | 930  |      |      |      | 935  |      |      |      |     | 940  |      |      |      |      |  |  |  |  |
| Gln  | Arg  | Lys  | Leu  | Ser  | Val  | Ala  | Ile  | Ala  | Phe  | Val | Gly  | Gly  | Ser  | Gln  | Val  |  |  |  |  |
| 945  |      |      |      |      | 950  |      |      |      | 955  |     |      |      |      |      | 960  |  |  |  |  |
| Val  | Ile  | Leu  | Asp  | Glu  | Pro  | Thr  | Ala  | Gly  | Val  | Asp | Pro  | Ala  | Ser  | Arg  | Arg  |  |  |  |  |
|      |      |      |      | 965  |      |      |      |      | 970  |     |      |      |      | 975  |      |  |  |  |  |
| Gly  | Ile  | Trp  | Glu  | Leu  | Leu  | Leu  | Lys  | Tyr  | Arg  | Glu | Gly  | Arg  | Thr  | Leu  | Ile  |  |  |  |  |
|      |      |      | 980  |      |      |      |      | 985  |      |     |      |      | 990  |      |      |  |  |  |  |
| Leu  | Ser  | Thr  | His  | His  | Leu  | Asp  | Glu  | Ala  | Glu  | Leu | Leu  | Gly  | Asp  | Arg  | Val  |  |  |  |  |
|      |      | 995  |      |      |      | 1000 |      |      |      |     |      | 1005 |      |      |      |  |  |  |  |
| Ala  | Val  | Val  | Ala  | Gly  | Gly  | Arg  | Leu  | Cys  | Cys  | Cys | Gly  | Ser  | Pro  | Leu  | Phe  |  |  |  |  |
|      | 1010 |      |      |      |      | 1015 |      |      |      |     | 1020 |      |      |      |      |  |  |  |  |
| Leu  | Arg  | Arg  | His  | Leu  | Gly  | Ser  | Gly  | Tyr  | Tyr  | Leu | Thr  | Leu  | Val  | Lys  | Ala  |  |  |  |  |
| 1025 |      |      |      |      | 1030 |      |      |      | 1035 |     |      |      |      |      | 1040 |  |  |  |  |
| Arg  | Leu  | Pro  | Leu  | Thr  | Thr  | Asn  | Glu  | Lys  | Ala  | Asp | Thr  | Asp  | Met  | Glu  | Gly  |  |  |  |  |
|      |      |      |      | 1045 |      |      |      | 1050 |      |     |      |      |      | 1055 |      |  |  |  |  |
| Ser  | Val  | Asp  | Thr  | Arg  | Gln  | Glu  | Lys  | Lys  | Asn  | Gly | Ser  | Gln  | Gly  | Ser  | Arg  |  |  |  |  |
|      |      | 1060 |      |      |      |      | 1065 |      |      |     |      |      | 1070 |      |      |  |  |  |  |
| Val  | Gly  | Thr  | Pro  | Gln  | Leu  | Leu  | Ala  | Leu  | Val  | Gln | His  | Trp  | Val  | Pro  | Gly  |  |  |  |  |
|      | 1075 |      |      |      |      | 1080 |      |      |      |     |      | 1085 |      |      |      |  |  |  |  |
| Ala  | Arg  | Leu  | Val  | Glu  | Glu  | Leu  | Pro  | His  | Glu  | Leu | Val  | Leu  | Val  | Leu  | Pro  |  |  |  |  |
|      | 1090 |      |      |      |      | 1095 |      |      |      |     | 1100 |      |      |      |      |  |  |  |  |
| Tyr  | Thr  | Gly  | Ala  | His  | Asp  | Gly  | Ser  | Phe  | Ala  | Thr | Leu  | Phe  | Arg  | Glu  | Leu  |  |  |  |  |
| 1105 |      |      |      | 1110 |      |      |      |      | 1115 |     |      |      |      | 1120 |      |  |  |  |  |
| Asp  | Thr  | Arg  | Leu  | Ala  | Glu  | Leu  | Arg  | Leu  | Thr  | Gly | Tyr  | Gly  | Ile  | Ser  | Asp  |  |  |  |  |
|      |      |      | 1125 |      |      |      |      | 1130 |      |     |      |      | 1135 |      |      |  |  |  |  |
| Thr  | Ser  | Leu  | Glu  | Glu  | Ile  | Phe  | Leu  | Lys  | Val  | Val | Glu  | Glu  | Cys  | Ala  | Ala  |  |  |  |  |
|      |      | 1140 |      |      |      | 1145 |      |      |      |     |      |      | 1150 |      |      |  |  |  |  |
| Asp  | Thr  | Asp  | Met  | Glu  | Asp  | Gly  | Ser  | Cys  | Gly  | Gln | His  | Leu  | Cys  | Thr  | Gly  |  |  |  |  |
|      | 1155 |      |      |      |      | 1160 |      |      |      |     | 1165 |      |      |      |      |  |  |  |  |
| Ile  | Ala  | Gly  | Leu  | Asp  | Val  | Thr  | Leu  | Arg  | Leu  | Lys | Met  | Pro  | Pro  | Gln  | Glu  |  |  |  |  |
|      | 1170 |      |      |      |      | 1175 |      |      |      |     | 1180 |      |      |      |      |  |  |  |  |
| Thr  | Ala  | Leu  | Glu  | Asn  | Gly  | Glu  | Pro  | Ala  | Gly  | Ser | Ala  | Pro  | Glu  | Thr  | Asp  |  |  |  |  |
| 1185 |      |      |      | 1190 |      |      |      |      | 1195 |     |      |      |      | 1200 |      |  |  |  |  |
| Gln  | Gly  | Ser  | Gly  | Pro  | Asp  | Ala  | Val  | Gly  | Arg  | Val | Gln  | Gly  | Trp  | Ala  | Leu  |  |  |  |  |
|      |      |      | 1205 |      |      |      |      | 1210 |      |     |      |      | 1215 |      |      |  |  |  |  |

Thr Arg Gln Gln Leu Gln Ala Leu Leu Lys Arg Phe Leu Leu Ala  
 1220 1225 1230  
 Arg Arg Ser Arg Arg Gly Leu Phe Ala Gln Ile Val Leu Pro Ala Leu  
 1235 1240 1245  
 Phe Val Gly Leu Ala Leu Val Phe Ser Leu Ile Val Pro Pro Phe Gly  
 1250 1255 1260  
 His Tyr Pro Ala Leu Arg Leu Ser Pro Thr Met Tyr Gly Ala Gln Val  
 1265 1270 1275 1280  
 Ser Phe Phe Ser Glu Asp Ala Pro Gly Asp Pro Gly Arg Ala Arg Leu  
 1285 1290 1295  
 Leu Glu Ala Leu Leu Gln Glu Ala Gly Leu Glu Glu Pro Pro Val Gln  
 1300 1305 1310  
 His Ser Ser His Arg Phe Ser Ala Pro Glu Val Pro Ala Glu Val Ala  
 1315 1320 1325  
 Lys Val Leu Ala Ser Gly Asn Trp Thr Pro Glu Ser Pro Ser Pro Ala  
 1330 1335 1340  
 Cys Gln Cys Ser Gln Pro Gly Ala Arg Arg Leu Pro Asp Cys Pro  
 1345 1350 1355 1360  
 Ala Ala Ala Gly Gly Pro Pro Pro Pro Gln Ala Val Thr Gly Ser Gly  
 1365 1370 1375  
 Glu Val Val Gln Asn Leu Thr Gly Arg Asn Leu Ser Asp Phe Leu Val  
 1380 1385 1390  
 Lys Thr Tyr Pro Arg Leu Val Arg Gln Gly Leu Lys Thr Lys Lys Trp  
 1395 1400 1405  
 Val Asn Glu Val Arg Tyr Gly Gly Phe Ser Leu Gly Gly Arg Asp Pro  
 1410 1415 1420  
 Gly Leu Pro Ser Gly Gln Glu Leu Gly Arg Ser Val Glu Glu Leu Trp  
 1425 1430 1435 1440  
 Ala Leu Leu Ser Pro Leu Pro Gly Gly Ala Leu Asp Arg Val Leu Lys  
 1445 1450 1455  
 Asn Leu Thr Ala Trp Ala His Ser Leu Asp Ala Gln Asp Ser Leu Lys  
 1460 1465 1470  
 Ile Trp Phe Asn Asn Lys Gly Trp His Ser Met Val Ala Phe Val Asn  
 1475 1480 1485  
 Arg Ala Ser Asn Ala Ile Leu Arg Ala His Leu Pro Gly Pro Ala  
 1490 1495 1500  
 Arg His Ala His Ser Ile Thr Thr Leu Asn His Pro Leu Asn Leu Thr  
 1505 1510 1515 1520

Lys Glu Gln Leu Ser Glu Ala Ala Leu Met Ala Ser Ser Val Asp Val  
 1525 1530 1535  
 Leu Val Ser Ile Cys Val Val Phe Ala Met Ser Phe Val Pro Ala Ser  
 1540 1545 1550  
 Phe Thr Leu Val Leu Ile Glu Glu Arg Val Thr Arg Ala Lys His Leu  
 1555 1560 1565  
 Gln Leu Met Gly Gly Leu Ser Pro Thr Leu Tyr Trp Leu Gly Asn Phe  
 1570 1575 1580  
 Leu Trp Asp Met Cys Asn Tyr Leu Val Pro Ala Cys Ile Val Val Leu  
 1585 1590 1595 1600  
 Ile Phe Leu Ala Phe Gln Gln Arg Ala Tyr Val Ala Pro Ala Asn Leu  
 1605 1610 1615  
 Pro Ala Leu Leu Leu Leu Leu Leu Tyr Gly Trp Ser Ile Thr Pro  
 1620 1625 1630  
 Leu Met Tyr Pro Ala Ser Phe Phe Phe Ser Val Pro Ser Thr Ala Tyr  
 1635 1640 1645  
 Val Val Leu Thr Cys Ile Asn Leu Phe Ile Gly Ile Asn Gly Ser Met  
 1650 1655 1660  
 Ala Thr Phe Val Leu Glu Leu Phe Ser Asp Gln Lys Leu Gln Glu Val  
 1665 1670 1675 1680  
 Ser Arg Ile Leu Lys Gln Val Phe Leu Ile Phe Pro His Phe Cys Leu  
 1685 1690 1695  
 Gly Arg Gly Leu Ile Asp Met Val Arg Asn Gln Ala Met Ala Asp Ala  
 1700 1705 1710  
 Phe Glu Arg Leu Gly Asp Arg Gln Phe Gln Ser Pro Leu Arg Trp Glu  
 1715 1720 1725  
 Val Val Gly Lys Asn Leu Leu Ala Met Val Ile Gln Gly Pro Leu Phe  
 1730 1735 1740  
 Leu Leu Phe Thr Leu Leu Leu Gln His Arg Ser Gln Leu Leu Pro Gln  
 1745 1750 1755 1760  
 Pro Arg Val Arg Ser Leu Pro Leu Leu Gly Glu Glu Asp Glu Asp Val  
 1765 1770 1775  
 Ala Arg Glu Arg Glu Arg Val Val Gln Gly Ala Thr Gln Gly Asp Val  
 1780 1785 1790  
 Leu Val Leu Arg Asn Leu Thr Lys Val Tyr Arg Gly Gln Arg Met Pro  
 1795 1800 1805  
 Ala Val Asp Arg Leu Cys Leu Gly Ile Pro Pro Gly Glu Cys Phe Gly  
 1810 1815 1820  
 Leu Leu Gly Val Asn Gly Ala Gly Lys Thr Ser Thr Phe Arg Met Val

|                                                                 |      |      |      |
|-----------------------------------------------------------------|------|------|------|
| 1825                                                            | 1830 | 1835 | 1840 |
| Thr Gly Asp Thr Leu Ala Ser Arg Gly Glu Ala Val Leu Ala Gly His | 1845 | 1850 | 1855 |
| Ser Val Ala Arg Glu Pro Ser Ala Ala His Leu Ser Met Gly Tyr Cys | 1860 | 1865 | 1870 |
| Pro Gln Ser Asp Ala Ile Phe Glu Leu Leu Thr Gly Arg Glu His Leu | 1875 | 1880 | 1885 |
| Glu Leu Leu Ala Arg Leu Arg Gly Val Pro Glu Ala Gln Val Ala Gln | 1890 | 1895 | 1900 |
| Thr Ala Gly Ser Gly Leu Ala Arg Leu Gly Leu Ser Trp Tyr Ala Asp | 1905 | 1910 | 1915 |
| Arg Pro Ala Gly Thr Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ala Thr | 1925 | 1930 | 1935 |
| Ala Leu Ala Leu Val Gly Asp Pro Ala Val Val Phe Leu Asp Glu Pro | 1940 | 1945 | 1950 |
| Thr Thr Gly Met Asp Pro Ser Ala Arg Arg Phe Leu Trp Asn Ser Leu | 1955 | 1960 | 1965 |
| Leu Ala Val Val Arg Glu Gly Arg Ser Val Met Leu Thr Ser His Ser | 1970 | 1975 | 1980 |
| Met Glu Glu Cys Glu Ala Leu Cys Ser Arg Leu Ala Ile Met Val Asn | 1985 | 1990 | 1995 |
| Gly Arg Phe Arg Cys Leu Gly Ser Pro Gln His Leu Lys Gly Arg Phe | 2005 | 2010 | 2015 |
| Ala Ala Gly His Thr Leu Thr Leu Arg Val Pro Ala Ala Arg Ser Gln | 2020 | 2025 | 2030 |
| Pro Ala Ala Ala Phe Val Ala Ala Glu Phe Pro Gly Ser Glu Leu Arg | 2035 | 2040 | 2045 |
| Glu Ala His Gly Gly Arg Leu Arg Phe Gln Leu Pro Pro Gly Gly Arg | 2050 | 2055 | 2060 |
| Cys Ala Leu Ala Arg Val Phe Gly Glu Leu Ala Val His Gly Ala Glu | 2065 | 2070 | 2075 |
| His Gly Val Glu Asp Phe Ser Val Ser Gln Thr Met Leu Glu Glu Val | 2085 | 2090 | 2095 |
| Phe Leu Tyr Phe Ser Lys Asp Gln Gly Lys Asp Glu Asp Thr Glu Glu | 2100 | 2105 | 2110 |
| Gln Lys Glu Ala Gly Val Gly Val Asp Pro Ala Pro Gly Leu Gln His | 2115 | 2120 | 2125 |
| Pro Lys Arg Val Ser Gln Phe Leu Asp Asp Pro Ser Thr Ala Glu Thr | 2130 | 2135 | 2140 |

Val Leu  
2145

<210> 6  
<211> 2100  
<212> PRT  
<213> Homo sapiens

<400> 6  
Pro Pro Leu Glu His His Glu Cys His Phe Pro Asn Lys Pro Leu Pro  
1 5 10 15  
Ser Ala Gly Thr Val Pro Trp Leu Gln Gly Leu Ile Cys Asn Val Asn  
20 25 30  
Asn Thr Cys Phe Pro Gln Leu Thr Pro Gly Glu Glu Pro Gly Arg Leu  
35 40 45  
Ser Asn Phe Asn Asp Ser Leu Val Ser Arg Leu Leu Ala Asp Ala Arg  
50 55 60  
Thr Val Leu Gly Gly Ala Ser Ala His Arg Thr Leu Ala Gly Leu Gly  
65 70 75 80  
Lys Leu Ile Ala Thr Leu Arg Ala Ala Arg Ser Thr Ala Gln Pro Gln  
85 90 95  
Pro Thr Lys Gln Ser Pro Leu Glu Pro Pro Met Leu Asp Val Ala Glu  
100 105 110  
Leu Leu Thr Ser Leu Leu Arg Thr Glu Ser Leu Gly Leu Ala Leu Gly  
115 120 125  
Gln Ala Gln Glu Pro Leu His Ser Leu Leu Glu Ala Ala Glu Asp Leu  
130 135 140  
Ala Gln Glu Leu Leu Ala Leu Arg Ser Leu Val Glu Leu Arg Ala Leu  
145 150 155 160  
Leu Gln Arg Pro Arg Gly Thr Ser Gly Pro Leu Glu Leu Leu Ser Glu  
165 170 175  
Ala Leu Cys Ser Val Arg Gly Pro Ser Ser Thr Val Gly Pro Ser Leu  
180 185 190  
Asn Trp Tyr Glu Ala Ser Asp Leu Met Glu Leu Val Gly Gln Glu Pro  
195 200 205  
Glu Ser Ala Leu Pro Asp Ser Ser Leu Ser Pro Ala Cys Ser Glu Leu  
210 215 220  
Ile Gly Ala Leu Asp Ser His Pro Leu Ser Arg Leu Leu Trp Arg Arg  
225 230 235 240  
Leu Lys Pro Leu Ile Leu Gly Lys Leu Leu Phe Ala Pro Asp Thr Pro  
245 250 255

Phe Thr Arg Lys Leu Met Ala Gln Val Asn Arg Thr Phe Glu Glu Leu  
 260 265 270  
 Thr Leu Leu Arg Asp Val Arg Glu Val Trp Glu Met Leu Gly Pro Arg  
 275 280 285  
 Ile Phe Thr Phe Met Asn Asp Ser Ser Asn Val Ala Met Leu Gln Arg  
 290 295 300  
 Leu Leu Gln Met Gln Asp Glu Gly Arg Arg Gln Pro Arg Pro Gly Gly  
 305 310 315 320  
 Arg Asp His Met Glu Ala Leu Arg Ser Phe Leu Asp Pro Gly Ser Gly  
 325 330 335  
 Gly Tyr Ser Trp Gln Asp Ala His Ala Asp Val Gly His Leu Val Gly  
 340 345 350  
 Thr Leu Gly Arg Val Thr Glu Cys Leu Ser Leu Asp Lys Leu Glu Ala  
 355 360 365  
 Ala Pro Ser Glu Ala Ala Leu Val Ser Arg Ala Leu Gln Leu Leu Ala  
 370 375 380  
 Glu His Arg Phe Trp Ala Gly Val Val Phe Leu Gly Pro Glu Asp Ser  
 385 390 395 400  
 Ser Asp Pro Thr Glu His Pro Thr Pro Asp Leu Gly Pro Gly His Val  
 405 410 415  
 Arg Ile Lys Ile Arg Met Asp Ile Asp Val Val Thr Arg Thr Asn Lys  
 420 425 430  
 Ile Arg Asp Arg Phe Trp Asp Pro Gly Pro Ala Ala Asp Pro Leu Thr  
 435 440 445  
 Asp Leu Arg Tyr Val Trp Gly Gly Phe Val Tyr Leu Gln Asp Leu Val  
 450 455 460  
 Glu Arg Ala Ala Val Arg Val Leu Ser Gly Ala Asn Pro Arg Ala Gly  
 465 470 475 480  
 Leu Tyr Leu Gln Gln Met Pro Tyr Pro Cys Tyr Val Asp Asp Val Phe  
 485 490 495  
 Leu Arg Val Leu Ser Arg Ser Leu Pro Leu Phe Leu Thr Leu Ala Trp  
 500 505 510  
 Ile Tyr Ser Val Thr Leu Thr Val Lys Ala Val Val Arg Glu Lys Glu  
 515 520 525  
 Thr Arg Leu Arg Asp Thr Met Arg Ala Met Gly Leu Ser Arg Ala Val  
 530 535 540  
 Leu Trp Leu Gly Trp Phe Leu Ser Cys Leu Gly Pro Phe Leu Leu Ser  
 545 550 555 560

Ala Ala Leu Leu Val Leu Val Leu Lys Leu Gly Asp Ile Leu Pro Tyr  
 565 570 575  
 Ser His Pro Gly Val Val Phe Leu Phe Leu Ala Ala Phe Ala Val Ala  
 580 585 590  
 Thr Val Thr Gln Ser Phe Leu Leu Ser Ala Phe Phe Ser Arg Ala Asn  
 595 600 605  
 Leu Ala Ala Ala Cys Gly Gly Leu Ala Tyr Phe Ser Leu Tyr Leu Pro  
 610 615 620  
 Tyr Val Leu Cys Val Ala Trp Arg Asp Arg Leu Pro Ala Gly Gly Arg  
 625 630 635 640  
 Val Ala Ala Ser Leu Leu Ser Pro Val Ala Phe Gly Phe Gly Cys Glu  
 645 650 655  
 Ser Leu Ala Leu Leu Glu Glu Gln Gly Glu Gly Ala Gln Trp His Asn  
 660 665 670  
 Val Gly Thr Arg Pro Thr Ala Asp Val Phe Ser Leu Ala Gln Val Ser  
 675 680 685  
 Gly Leu Leu Leu Leu Asp Ala Ala Leu Tyr Gly Leu Ala Thr Trp Tyr  
 690 695 700  
 Leu Glu Ala Val Cys Pro Gly Gln Tyr Gly Ile Pro Glu Pro Trp Asn  
 705 710 715 720  
 Phe Pro Phe Arg Arg Ser Tyr Trp Cys Gly Pro Arg Pro Pro Lys Ser  
 725 730 735  
 Pro Ala Pro Cys Pro Thr Pro Leu Asp Pro Lys Val Leu Val Glu Glu  
 740 745 750  
 Ala Pro Pro Gly Leu Ser Pro Gly Val Ser Val Arg Ser Leu Glu Lys  
 755 760 765  
 Arg Phe Pro Gly Ser Pro Gln Pro Ala Leu Arg Gly Leu Ser Leu Asp  
 770 775 780  
 Phe Tyr Gln Gly His Ile Thr Ala Phe Leu Gly His Asn Gly Ala Gly  
 785 790 795 800  
 Lys Thr Thr Thr Leu Ser Ile Leu Ser Gly Leu Phe Pro Pro Ser Gly  
 805 810 815  
 Gly Ser Ala Phe Ile Leu Gly His Asp Val Arg Ser Ser Met Ala Ala  
 820 825 830  
 Ile Arg Pro His Leu Gly Val Cys Pro Gln Tyr Asn Val Leu Phe Asp  
 835 840 845  
 Met Leu Thr Val Asp Glu His Val Trp Phe Tyr Gly Arg Leu Lys Gly  
 850 855 860  
 Leu Ser Ala Ala Val Val Gly Pro Glu Gln Asp Arg Leu Leu Gln Asp

|                 |                             |                             |             |     |  |      |
|-----------------|-----------------------------|-----------------------------|-------------|-----|--|------|
| 865             |                             | 870                         |             | 875 |  | 880  |
| Val Gly Leu Val | Ser Lys Gln Ser Val         | Gln Thr Arg His Leu Ser Gly |             |     |  |      |
|                 | 885                         | 890                         |             |     |  | 895  |
| Gly Met Gln Arg | Lys Leu Ser Val             | Ala Ile Ala Phe Val         | Gly Gly Ser |     |  |      |
|                 | 900                         | 905                         | 910         |     |  |      |
| Gln Val Val Ile | Leu Asp Glu Pro Thr         | Ala Gly Val Asp             | Pro Ala Ser |     |  |      |
|                 | 915                         | 920                         | 925         |     |  |      |
| Arg Arg Gly Ile | Trp Glu Leu Leu Lys Tyr     | Arg Glu Gly Arg Thr         |             |     |  |      |
|                 | 930                         | 935                         | 940         |     |  |      |
| Leu Ile Leu Ser | Thr His His Leu Asp         | Glu Ala Glu Leu Leu Gly     | Asp         |     |  |      |
|                 | 945                         | 950                         | 955         |     |  | 960  |
| Arg Val Ala Val | Val Ala Gly Gly Arg         | Leu Cys Cys Cys Gly         | Ser Pro     |     |  |      |
|                 | 965                         | 970                         | 975         |     |  |      |
| Leu Phe Leu Arg | Arg His Leu Gly Ser         | Gly Tyr Tyr Leu Thr         | Leu Val     |     |  |      |
|                 | 980                         | 985                         | 990         |     |  |      |
| Lys Ala Arg Leu | Pro Leu Thr Thr             | Asn Glu Lys Ala Asp         | Thr Asp Met |     |  |      |
|                 | 995                         | 1000                        | 1005        |     |  |      |
| Glu Gly Ser Val | Asp Thr Arg Gln Glu Lys     | Lys Asn Gly Ser Gln Gly     |             |     |  |      |
|                 | 1010                        | 1015                        | 1020        |     |  |      |
| Ser Arg Val Gly | Thr Pro Gln Leu Leu Ala Leu | Val Gln His Trp Val         |             |     |  |      |
|                 | 1025                        | 1030                        | 1035        |     |  | 1040 |
| Pro Gly Ala Arg | Leu Val Glu Glu Leu Pro     | His Glu Leu Val Leu Val     |             |     |  |      |
|                 | 1045                        | 1050                        | 1055        |     |  |      |
| Leu Pro Tyr Thr | Gly Ala His Asp Gly         | Ser Phe Ala Thr Leu Phe Arg |             |     |  |      |
|                 | 1060                        | 1065                        | 1070        |     |  |      |
| Glu Leu Asp Thr | Arg Leu Ala Glu Leu Arg     | Leu Thr Gly Tyr Gly Ile     |             |     |  |      |
|                 | 1075                        | 1080                        | 1085        |     |  |      |
| Ser Asp Thr Ser | Leu Glu Glu Ile Phe Leu     | Lys Val Val Glu Glu Cys     |             |     |  |      |
|                 | 1090                        | 1095                        | 1100        |     |  |      |
| Ala Ala Asp Thr | Asp Met Glu Asp Gly Ser     | Cys Gly Gln His Leu Cys     |             |     |  |      |
|                 | 1105                        | 1110                        | 1115        |     |  | 1120 |
| Thr Gly Ile Ala | Gly Leu Asp Val Thr Leu     | Arg Leu Lys Met Pro Pro     |             |     |  |      |
|                 | 1125                        | 1130                        | 1135        |     |  |      |
| Gln Glu Thr Ala | Leu Glu Asn Gly Glu Pro     | Ala Gly Ser Ala Pro Glu     |             |     |  |      |
|                 | 1140                        | 1145                        | 1150        |     |  |      |
| Thr Asp Gln Gly | Ser Gly Pro Asp Ala Val     | Gly Arg Val Gln Gly Trp     |             |     |  |      |
|                 | 1155                        | 1160                        | 1165        |     |  |      |
| Ala Leu Thr Arg | Gln Gln Leu Gln Ala Leu     | Leu Leu Lys Arg Phe Leu     |             |     |  |      |
|                 | 1170                        | 1175                        | 1180        |     |  |      |

Leu Ala Arg Arg Ser Arg Arg Gly Leu Phe Ala Gln Ile Val Leu Pro  
 1185 1190 1195 1200  
 Ala Leu Phe Val Gly Leu Ala Leu Val Phe Ser Leu Ile Val Pro Pro  
 1205 1210 1215  
 Phe Gly His Tyr Pro Ala Leu Arg Leu Ser Pro Thr Met Tyr Gly Ala  
 1220 1225 1230  
 Gln Val Ser Phe Phe Ser Glu Asp Ala Pro Gly Asp Pro Gly Arg Ala  
 1235 1240 1245  
 Arg Leu Leu Glu Ala Leu Leu Gln Glu Ala Gly Leu Glu Glu Pro Pro  
 1250 1255 1260  
 Val Gln His Ser Ser His Arg Phe Ser Ala Pro Glu Val Pro Ala Glu  
 1265 1270 1275 1280  
 Val Ala Lys Val Leu Ala Ser Gly Asn Trp Thr Pro Glu Ser Pro Ser  
 1285 1290 1295  
 Pro Ala Cys Gln Cys Ser Gln Pro Gly Ala Arg Arg Leu Leu Pro Asp  
 1300 1305 1310  
 Cys Pro Ala Ala Ala Gly Gly Pro Pro Pro Gln Ala Val Thr Gly  
 1315 1320 1325  
 Ser Gly Glu Val Val Gln Asn Leu Thr Gly Arg Asn Leu Ser Asp Phe  
 1330 1335 1340  
 Leu Val Lys Thr Tyr Pro Arg Leu Val Arg Gln Gly Leu Lys Thr Lys  
 1345 1350 1355 1360  
 Lys Trp Val Asn Glu Val Arg Tyr Gly Gly Phe Ser Leu Gly Gly Arg  
 1365 1370 1375  
 Asp Pro Gly Leu Pro Ser Gly Gln Glu Leu Gly Arg Ser Val Glu Glu  
 1380 1385 1390  
 Leu Trp Ala Leu Leu Ser Pro Leu Pro Gly Gly Ala Leu Asp Arg Val  
 1395 1400 1405  
 Leu Lys Asn Leu Thr Ala Trp Ala His Ser Leu Asp Ala Gln Asp Ser  
 1410 1415 1420  
 Leu Lys Ile Trp Phe Asn Asn Lys Gly Trp His Ser Met Val Ala Phe  
 1425 1430 1435 1440  
 Val Asn Arg Ala Ser Asn Ala Ile Leu Arg Ala His Leu Pro Pro Gly  
 1445 1450 1455  
 Pro Ala Arg His Ala His Ser Ile Thr Thr Leu Asn His Pro Leu Asn  
 1460 1465 1470  
 Leu Thr Lys Glu Gln Leu Ser Glu Ala Ala Leu Met Ala Ser Ser Val  
 1475 1480 1485

Asp Val Leu Val Ser Ile Cys Val Val Phe Ala Met Ser Phe Val Pro  
 1490 1495 1500  
 Ala Ser Phe Thr Leu Val Leu Ile Glu Glu Arg Val Thr Arg Ala Lys  
 1505 1510 1515 1520  
 His Leu Gln Leu Met Gly Gly Leu Ser Pro Thr Leu Tyr Trp Leu Gly  
 1525 1530 1535  
 Asn Phe Leu Trp Asp Met Cys Asn Tyr Leu Val Pro Ala Cys Ile Val  
 1540 1545 1550  
 Val Leu Ile Phe Leu Ala Phe Gln Gln Arg Ala Tyr Val Ala Pro Ala  
 1555 1560 1565  
 Asn Leu Pro Ala Leu Leu Leu Leu Leu Leu Tyr Gly Trp Ser Ile  
 1570 1575 1580  
 Thr Pro Leu Met Tyr Pro Ala Ser Phe Phe Phe Ser Val Pro Ser Thr  
 1585 1590 1595 1600  
 Ala Tyr Val Val Leu Thr Cys Ile Asn Leu Phe Ile Gly Ile Asn Gly  
 1605 1610 1615  
 Ser Met Ala Thr Phe Val Leu Glu Leu Phe Ser Asp Gln Lys Leu Gln  
 1620 1625 1630  
 Glu Val Ser Arg Ile Leu Lys Gln Val Phe Leu Ile Phe Pro His Phe  
 1635 1640 1645  
 Cys Leu Gly Arg Gly Leu Ile Asp Met Val Arg Asn Gln Ala Met Ala  
 1650 1655 1660  
 Asp Ala Phe Glu Arg Leu Gly Asp Arg Gln Phe Gln Ser Pro Leu Arg  
 1665 1670 1675 1680  
 Trp Glu Val Val Gly Lys Asn Leu Leu Ala Met Val Ile Gln Gly Pro  
 1685 1690 1695  
 Leu Phe Leu Leu Phe Thr Leu Leu Leu Gln His Arg Ser Gln Leu Leu  
 1700 1705 1710  
 Pro Gln Pro Arg Val Arg Ser Leu Pro Leu Leu Gly Glu Glu Asp Glu  
 1715 1720 1725  
 Asp Val Ala Arg Glu Arg Glu Arg Val Val Gln Gly Ala Thr Gln Gly  
 1730 1735 1740  
 Asp Val Leu Val Leu Arg Asn Leu Thr Lys Val Tyr Arg Gly Gln Arg  
 1745 1750 1755 1760  
 Met Pro Ala Val Asp Arg Leu Cys Leu Gly Ile Pro Pro Gly Glu Cys  
 1765 1770 1775  
 Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Thr Ser Thr Phe Arg  
 1780 1785 1790  
 Met Val Thr Gly Asp Thr Leu Ala Ser Arg Gly Glu Ala Val Leu Ala

| 1795                                                                                   | 1800 | 1805 |
|----------------------------------------------------------------------------------------|------|------|
| Gly His Ser Val Ala Arg Glu Pro Ser Ala Ala His Leu Ser Met Gly<br>1810 1815 1820      |      |      |
| Tyr Cys Pro Gln Ser Asp Ala Ile Phe Glu Leu Leu Thr Gly Arg Glu<br>1825 1830 1835 1840 |      |      |
| His Leu Glu Leu Leu Ala Arg Leu Arg Gly Val Pro Glu Ala Gln Val<br>1845 1850 1855      |      |      |
| Ala Gln Thr Ala Gly Ser Gly Leu Ala Arg Leu Gly Leu Ser Trp Tyr<br>1860 1865 1870      |      |      |
| Ala Asp Arg Pro Ala Gly Thr Tyr Ser Gly Gly Asn Lys Arg Lys Leu<br>1875 1880 1885      |      |      |
| Ala Thr Ala Leu Ala Leu Val Gly Asp Pro Ala Val Val Phe Leu Asp<br>1890 1895 1900      |      |      |
| Glu Pro Thr Thr Gly Met Asp Pro Ser Ala Arg Arg Phe Leu Trp Asn<br>1905 1910 1915 1920 |      |      |
| Ser Leu Leu Ala Val Val Arg Glu Gly Arg Ser Val Met Leu Thr Ser<br>1925 1930 1935      |      |      |
| His Ser Met Glu Glu Cys Glu Ala Leu Cys Ser Arg Leu Ala Ile Met<br>1940 1945 1950      |      |      |
| Val Asn Gly Arg Phe Arg Cys Leu Gly Ser Pro Gln His Leu Lys Gly<br>1955 1960 1965      |      |      |
| Arg Phe Ala Ala Gly His Thr Leu Thr Leu Arg Val Pro Ala Ala Arg<br>1970 1975 1980      |      |      |
| Ser Gln Pro Ala Ala Ala Phe Val Ala Ala Glu Phe Pro Gly Ser Glu<br>1985 1990 1995 2000 |      |      |
| Leu Arg Glu Ala His Gly Gly Arg Leu Arg Phe Gln Leu Pro Pro Gly<br>2005 2010 2015      |      |      |
| Gly Arg Cys Ala Leu Ala Arg Val Phe Gly Glu Leu Ala Val His Gly<br>2020 2025 2030      |      |      |
| Ala Glu His Gly Val Glu Asp Phe Ser Val Ser Gln Thr Met Leu Glu<br>2035 2040 2045      |      |      |
| Glu Val Phe Leu Tyr Phe Ser Lys Asp Gln Gly Lys Asp Glu Asp Thr<br>2050 2055 2060      |      |      |
| Glu Glu Gln Lys Glu Ala Gly Val Gly Val Asp Pro Ala Pro Gly Leu<br>2065 2070 2075 2080 |      |      |
| Gln His Pro Lys Arg Val Ser Gln Phe Leu Asp Asp Pro Ser Thr Ala<br>2085 2090 2095      |      |      |
| Glu Thr Val Leu<br>2100                                                                |      |      |

<210> 7  
 <211> 4653  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> CDS  
 <222> (1)..(4653)

<400> 7  
 atg gtc tgc ctg gga act ggc cag agc gct gga ccc cta gtg agt gtt 48  
 Met Val Cys Leu Gly Thr Gly Gln Ser Ala Gly Pro Leu Val Ser Val  
     1                    5                    10                    15

caa aat cat tgt ccc cct tgt ggt ctt tct ccc cag gaa tcc ctg ggg 96  
 Gln Asn His Cys Pro Pro Cys Gly Leu Ser Pro Gln Glu Ser Leu Gly  
                     20                    25                    30

ttg gca ctg ggc caa gcc cag gag ccc ttg cac agc ttg ttg gag gcc 144  
 Leu Ala Leu Gly Gln Ala Gln Glu Pro Leu His Ser Leu Leu Glu Ala  
                     35                    40                    45

gct ggg gac ctg gcc cag gag ctc ctg gcg ctg cgc agc ctg gtg gag 192  
 Ala Gly Asp Leu Ala Gln Glu Leu Leu Ala Leu Arg Ser Leu Val Glu  
                     50                    55                    60

ctt cgg gca ctg ctg cag aga ccc cga ggg acc agc ggc ccc ctg gag 240  
 Leu Arg Ala Leu Leu Gln Arg Pro Arg Gly Thr Ser Gly Pro Leu Glu  
     65                    70                    75                    80

ttg ctg tca gag gcc ctc tgc agt gtc agg gga cct agc agc aca gtg 288  
 Leu Leu Ser Glu Ala Leu Cys Ser Val Arg Gly Pro Ser Ser Thr Val  
                     85                    90                    95

ggc ccc tcc ctc aac tgg tac gag gct agt gac ctg atg gag ctg gtg 336  
 Gly Pro Ser Leu Asn Trp Tyr Glu Ala Ser Asp Leu Met Glu Leu Val  
                     100                    105                    110

ggg cag gag cca gaa tcc gcc ctg cca gac agc agc ctg agc ccc gcc 384  
 Gly Gln Glu Pro Glu Ser Ala Leu Pro Asp Ser Ser Leu Ser Pro Ala  
                     115                    120                    125

tgc tcg gag ctg att gga gcc ctg gac agc cac ccg ctg tcc cgc ctg 432  
 Cys Ser Glu Leu Ile Gly Ala Leu Asp Ser His Pro Leu Ser Arg Leu  
     130                    135                    140

ctc tgg aga cgc ctg aag cct ctg atc ctc ggg aag cta ctc ttt gca 480  
 Leu Trp Arg Arg Leu Lys Pro Leu Ile Leu Gly Lys Leu Leu Phe Ala  
     145                    150                    155                    160

cca gat aca cct ttt acc cgg aag ctc atg gcc cag gtg aac cgg acc 528  
 Pro Asp Thr Pro Phe Thr Arg Lys Leu Met Ala Gln Val Asn Arg Thr  
                     165                    170                    175

ttc gag gag ctc acc ctg ctg agg gat gtc cgg gag gtg tgg gag atg 576  
 Phe Glu Glu Leu Thr Leu Leu Arg Asp Val Arg Glu Val Trp Glu Met

| 180                                                             |      |  |  |  |  |  |  |  |  | 185 |  |  |  |  | 190 |  |  |  |  |  |
|-----------------------------------------------------------------|------|--|--|--|--|--|--|--|--|-----|--|--|--|--|-----|--|--|--|--|--|
| ctg gga ccc cgg atc ttc acc ttc atg aac gac agt tcc aat gtg gcc | 624  |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| Leu Gly Pro Arg Ile Phe Thr Phe Met Asn Asp Ser Ser Asn Val Ala |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 195 200 205                                                     |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| atg ctg cag cgg ctc ctg cag atg cag gat gaa gga aga agg cag ccc | 672  |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| Met Leu Gln Arg Leu Leu Gln Met Gln Asp Glu Gly Arg Arg Gln Pro |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 210 215 220                                                     |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| aga cct gga ggc cgg gac cac atg gag gcc ctg cga tcc ttt ctg gac | 720  |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| Arg Pro Gly Gly Arg Asp His Met Glu Ala Leu Arg Ser Phe Leu Asp |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 225 230 235 240                                                 |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| cct ggg agc ggt ggc tac agc tgg cag gac gca cac gct gat gtg ggg | 768  |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| Pro Gly Ser Gly Gly Tyr Ser Trp Gln Asp Ala His Ala Asp Val Gly |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 245 250 255                                                     |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| cac ctg gtg ggc acg ctg ggc cga gtg acg gag tgc ctg tcc ttg gac | 816  |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| His Leu Val Gly Thr Leu Gly Arg Val Thr Glu Cys Leu Ser Leu Asp |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 260 265 270                                                     |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| aag ctg gag gcg gca ccc tca gag gca gcc ctg gtg tcg cgg gcc ctg | 864  |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| Lys Leu Glu Ala Ala Pro Ser Glu Ala Ala Leu Val Ser Arg Ala Leu |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 275 280 285                                                     |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| caa ctg ctc gcg gaa cat cga ttc tgg gcc ggc gtc gtc ttc ttg gga | 912  |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| Gln Leu Leu Ala Glu His Arg Phe Trp Ala Gly Val Val Phe Leu Gly |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 290 295 300                                                     |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| cct gag gac tct tca gac ccc aca gag cac cca acc cca gac ctg ggc | 960  |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| Pro Glu Asp Ser Ser Asp Pro Thr Glu His Pro Thr Pro Asp Leu Gly |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 305 310 315 320                                                 |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| ccc ggc cac gtg cgc atc aaa atc cgc atg gac att gac gtg gtc acg | 1008 |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| Pro Gly His Val Arg Ile Lys Ile Arg Met Asp Ile Asp Val Val Thr |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 325 330 335                                                     |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| agg acc aat aag atc agg gac agg ttt tgg gac cct ggc cca gcc gcg | 1056 |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| Arg Thr Asn Lys Ile Arg Asp Arg Phe Trp Asp Pro Gly Pro Ala Ala |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 340 345 350                                                     |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| gac ccc ctg acc gac ctg cgc tac gtg tgg ggc ggc ttc gtg tac ctg | 1104 |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| Asp Pro Leu Thr Asp Leu Arg Tyr Val Trp Gly Gly Phe Val Tyr Leu |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 355 360 365                                                     |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| caa gac ctg gtg gag cgt gca gcc gtc cgc gtg ctc agc ggc gcc aac | 1152 |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| Gln Asp Leu Val Glu Arg Ala Ala Val Arg Val Leu Ser Gly Ala Asn |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 370 375 380                                                     |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| ccc cgg gcc ggc ctc tac ctg cag cag atg ccc tat ccg tgc tat gtg | 1200 |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| Pro Arg Ala Gly Leu Tyr Leu Gln Gln Met Pro Tyr Pro Cys Tyr Val |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 385 390 395 400                                                 |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| gac gac gtg ttc ctg cgt gtg ctg agc cgg tcg ctg ccg ctc ttc ctg | 1248 |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| Asp Asp Val Phe Leu Arg Val Leu Ser Arg Ser Leu Pro Leu Phe Leu |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |
| 405 410 415                                                     |      |  |  |  |  |  |  |  |  |     |  |  |  |  |     |  |  |  |  |  |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| acg | ctg | gcc | tgg | atc | tac | tcc | gtg | aca | ctg | aca | gtg | aag | gcc | gtg | gtg | 1296 |
| Thr | Leu | Ala | Trp | Ile | Tyr | Ser | Val | Thr | Leu | Thr | Val | Lys | Ala | Val | Val |      |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     | 430 |     |     |      |
| cgg | gag | aag | gag | acg | cgg | ctg | cgg | gac | acc | atg | cgc | gcc | atg | ggg | ctc | 1344 |
| Arg | Glu | Lys | Glu | Thr | Arg | Leu | Arg | Asp | Thr | Met | Arg | Ala | Met | Gly | Leu |      |
|     |     | 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |      |
| agc | cgc | gcg | gtg | ctc | tgg | cta | ggc | tgg | ttc | ctc | agc | tgc | ctc | ggg | ccc | 1392 |
| Ser | Arg | Ala | Val | Leu | Trp | Leu | Gly | Trp | Phe | Leu | Ser | Cys | Leu | Gly | Pro |      |
|     |     | 450 |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |      |
| ttc | ctg | ctc | agc | gcc | gcg | ctg | ctg | gtt | ctg | gtg | ctc | aag | ctg | ggg | gac | 1440 |
| Phe | Leu | Leu | Ser | Ala | Ala | Leu | Leu | Val | Leu | Val | Leu | Lys | Leu | Gly | Asp |      |
| 465 |     |     |     |     | 470 |     |     |     | 475 |     |     |     |     |     | 480 |      |
| atc | ctc | ccc | tac | agc | cac | ccg | ggc | gtg | gtc | ttc | ctg | ttc | ttg | gca | gcc | 1488 |
| Ile | Leu | Pro | Tyr | Ser | His | Pro | Gly | Val | Val | Phe | Leu | Phe | Leu | Ala | Ala |      |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     | 495 |     |      |
| ttc | gcg | gtg | gcc | acg | gtg | acc | cag | agc | ttc | ctg | ctc | agc | gcc | ttc | ttc | 1536 |
| Phe | Ala | Val | Ala | Thr | Val | Thr | Gln | Ser | Phe | Leu | Leu | Ser | Ala | Phe | Phe |      |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |     |     |      |
| tcc | cgc | gcc | aac | ctg | gct | gcg | gcc | tgc | ggc | ggc | ctg | gcc | tac | ttc | tcc | 1584 |
| Ser | Arg | Ala | Asn | Leu | Ala | Ala | Ala | Cys | Gly | Gly | Leu | Ala | Tyr | Phe | Ser |      |
|     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |     |      |
| ctc | tac | ctg | ccc | tac | gtg | ctg | tgt | gtg | gct | tgg | cgg | gac | cgg | ctg | ccc | 1632 |
| Leu | Tyr | Leu | Pro | Tyr | Val | Leu | Cys | Val | Ala | Trp | Arg | Asp | Arg | Leu | Pro |      |
|     |     | 530 |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |      |
| gcg | ggt | ggc | cgc | gtg | gcc | gcg | agc | ctg | ctg | tgc | ccc | gtg | gcc | ttc | ggc | 1680 |
| Ala | Gly | Gly | Arg | Val | Ala | Ala | Ser | Leu | Leu | Ser | Pro | Val | Ala | Phe | Gly |      |
| 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |      |
| ttc | ggc | tgc | gag | agc | ctg | gct | ctg | ctg | gag | gag | cag | ggc | gag | ggc | gcg | 1728 |
| Phe | Gly | Cys | Glu | Ser | Leu | Ala | Leu | Leu | Glu | Glu | Gln | Gly | Glu | Gly | Ala |      |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |      |
| cag | tgg | cac | aac | gtg | ggc | acc | cgg | cct | acg | gca | gac | gtc | ttc | agc | ctg | 1776 |
| Gln | Trp | His | Asn | Val | Gly | Thr | Arg | Pro | Thr | Ala | Asp | Val | Phe | Ser | Leu |      |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |      |
| gcc | cag | gtc | tct | ggc | ctt | ctg | ctg | ctg | gac | gcg | gcg | ctc | tac | ggc | ctc | 1824 |
| Ala | Gln | Val | Ser | Gly | Leu | Leu | Leu | Leu | Asp | Ala | Ala | Leu | Tyr | Gly | Leu |      |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |      |
| gcc | acc | tgg | tac | ctg | gaa | gct | gtg | tgc | cca | ggc | cag | tac | ggg | atc | cct | 1872 |
| Ala | Thr | Trp | Tyr | Leu | Glu | Ala | Val | Cys | Pro | Gly | Gln | Tyr | Gly | Ile | Pro |      |
|     |     | 610 |     |     |     | 615 |     |     |     |     | 620 |     |     |     |     |      |
| gaa | cca | tgg | aat | ttt | cct | ttt | cgg | agg | agc | tac | tgg | tgc | gga | cct | cgg | 1920 |
| Glu | Pro | Trp | Asn | Phe | Pro | Phe | Arg | Arg | Ser | Tyr | Trp | Cys | Gly | Pro | Arg |      |
| 625 |     |     |     |     | 630 |     |     |     | 635 |     |     |     |     |     | 640 |      |
| ccc | ccc | aag | agt | cca | gcc | cct | tgc | ccc | acc | ccg | ctg | gac | cca | aag | gtg | 1968 |
| Pro | Pro | Lys | Ser | Pro | Ala | Pro | Cys | Pro | Thr | Pro | Leu | Asp | Pro | Lys | Val |      |

| 645 |     |     |     |     |     |     |     | 650 |     |     |     |     |     |     |     | 655  |  |  |  |  |  |  |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|--|--|--|--|--|--|--|
| ctg | gta | gaa | gag | gca | ccg | ccc | ggc | ctg | agt | cct | ggc | gta | tcc | gtt | cgc | 2016 |  |  |  |  |  |  |  |
| Leu | Val | Glu | Glu | Ala | Pro | Pro | Gly | Leu | Ser | Pro | Gly | Val | Ser | Val | Arg |      |  |  |  |  |  |  |  |
| 660 |     |     |     |     |     |     |     | 665 |     |     |     |     |     |     |     | 670  |  |  |  |  |  |  |  |
| agc | ctg | gag | aag | cgc | ttt | cct | gga | agc | ccg | cag | cca | gcc | ctg | cgg | ggg | 2064 |  |  |  |  |  |  |  |
| Ser | Leu | Glu | Lys | Arg | Phe | Pro | Gly | Ser | Pro | Gln | Pro | Ala | Leu | Arg | Gly |      |  |  |  |  |  |  |  |
| 675 |     |     |     |     |     |     |     | 680 |     |     |     |     |     |     |     | 685  |  |  |  |  |  |  |  |
| ctc | agc | ctg | gac | ttc | tac | cag | ggc | cac | atc | acc | gcc | ttc | ctg | ggc | cac | 2112 |  |  |  |  |  |  |  |
| Leu | Ser | Leu | Asp | Phe | Tyr | Gln | Gly | His | Ile | Thr | Ala | Phe | Leu | Gly | His |      |  |  |  |  |  |  |  |
| 690 |     |     |     |     |     |     |     | 695 |     |     |     |     |     |     |     | 700  |  |  |  |  |  |  |  |
| aac | ggg | gcc | ggc | aag | acc | acc | acc | ctg | tcc | atc | ttg | agt | ggc | ctc | ttc | 2160 |  |  |  |  |  |  |  |
| Asn | Gly | Ala | Gly | Lys | Thr | Thr | Thr | Leu | Ser | Ile | Leu | Ser | Gly | Leu | Phe |      |  |  |  |  |  |  |  |
| 705 |     |     |     |     |     |     |     | 710 |     |     |     |     |     |     |     | 715  |  |  |  |  |  |  |  |
| cca | ccc | agt | ggt | ggc | tct | gcc | ttc | atc | ctg | ggc | cac | gac | gtc | cgc | tcc | 2208 |  |  |  |  |  |  |  |
| Pro | Pro | Ser | Gly | Gly | Ser | Ala | Phe | Ile | Leu | Gly | His | Asp | Val | Arg | Ser |      |  |  |  |  |  |  |  |
| 725 |     |     |     |     |     |     |     | 730 |     |     |     |     |     |     |     | 735  |  |  |  |  |  |  |  |
| agc | atg | gcc | gcc | atc | cgg | ccc | cac | ctg | ggc | gtc | tgt | cct | cag | tac | aac | 2256 |  |  |  |  |  |  |  |
| Ser | Met | Ala | Ala | Ile | Arg | Pro | His | Leu | Gly | Val | Cys | Pro | Gln | Tyr | Asn |      |  |  |  |  |  |  |  |
| 740 |     |     |     |     |     |     |     | 745 |     |     |     |     |     |     |     | 750  |  |  |  |  |  |  |  |
| gtg | ctg | ttt | gac | atg | ctg | acc | gtg | gac | gag | cac | gtc | tgg | ttc | tat | ggg | 2304 |  |  |  |  |  |  |  |
| Val | Leu | Phe | Asp | Met | Leu | Thr | Val | Asp | Glu | His | Val | Trp | Phe | Tyr | Gly |      |  |  |  |  |  |  |  |
| 755 |     |     |     |     |     |     |     | 760 |     |     |     |     |     |     |     | 765  |  |  |  |  |  |  |  |
| cgg | ctg | aag | ggt | ctg | agt | gcc | gct | gta | gtg | ggc | ccc | gag | cag | gac | cgt | 2352 |  |  |  |  |  |  |  |
| Arg | Leu | Lys | Gly | Leu | Ser | Ala | Ala | Val | Val | Gly | Pro | Glu | Gln | Asp | Arg |      |  |  |  |  |  |  |  |
| 770 |     |     |     |     |     |     |     | 775 |     |     |     |     |     |     |     | 780  |  |  |  |  |  |  |  |
| ctg | ctg | cag | gat | gtg | ggg | ctg | gtc | tcc | aag | cag | agt | gtg | cag | act | cgc | 2400 |  |  |  |  |  |  |  |
| Leu | Leu | Gln | Asp | Val | Gly | Leu | Val | Ser | Lys | Gln | Ser | Val | Gln | Thr | Arg |      |  |  |  |  |  |  |  |
| 785 |     |     |     |     |     |     |     | 790 |     |     |     |     |     |     |     | 800  |  |  |  |  |  |  |  |
| cac | ctc | tct | ggt | ggg | atg | caa | cgg | aag | ctg | tcc | gtg | gcc | att | gcc | ttt | 2448 |  |  |  |  |  |  |  |
| His | Leu | Ser | Gly | Gly | Met | Gln | Arg | Lys | Leu | Ser | Val | Ala | Ile | Ala | Phe |      |  |  |  |  |  |  |  |
| 805 |     |     |     |     |     |     |     | 810 |     |     |     |     |     |     |     | 815  |  |  |  |  |  |  |  |
| gtg | ggc | ggc | tcc | caa | gtt | gtt | atc | ctg | gac | gag | cct | acg | gct | ggc | gtg | 2496 |  |  |  |  |  |  |  |
| Val | Gly | Gly | Ser | Gln | Val | Val | Ile | Leu | Asp | Glu | Pro | Thr | Ala | Gly | Val |      |  |  |  |  |  |  |  |
| 820 |     |     |     |     |     |     |     | 825 |     |     |     |     |     |     |     | 830  |  |  |  |  |  |  |  |
| gat | cct | gct | tcc | cgc | cgc | ggt | att | tgg | gag | ctg | ctg | ctc | aaa | tac | cga | 2544 |  |  |  |  |  |  |  |
| Asp | Pro | Ala | Ser | Arg | Arg | Gly | Ile | Trp | Glu | Leu | Leu | Leu | Lys | Tyr | Arg |      |  |  |  |  |  |  |  |
| 835 |     |     |     |     |     |     |     | 840 |     |     |     |     |     |     |     | 845  |  |  |  |  |  |  |  |
| gaa | ggt | cgc | acg | ctg | atc | ctc | tcc | acc | cac | cac | ctg | gat | gag | gca | gag | 2592 |  |  |  |  |  |  |  |
| Glu | Gly | Arg | Thr | Leu | Ile | Leu | Ser | Thr | His | His | Leu | Asp | Glu | Ala | Glu |      |  |  |  |  |  |  |  |
| 850 |     |     |     |     |     |     |     | 855 |     |     |     |     |     |     |     | 860  |  |  |  |  |  |  |  |
| ctg | ctg | gga | gac | cgt | gtg | gct | gtg | gtg | gca | ggt | ggc | cgc | ttg | tgc | tgc | 2640 |  |  |  |  |  |  |  |
| Leu | Leu | Gly | Asp | Arg | Val | Ala | Val | Val | Ala | Gly | Gly | Arg | Leu | Cys | Cys |      |  |  |  |  |  |  |  |
| 865 |     |     |     |     |     |     |     | 870 |     |     |     |     |     |     |     | 875  |  |  |  |  |  |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |  |  |  |  |  |  |  |

|                                                                 |      |
|-----------------------------------------------------------------|------|
| tgt ggc tcc cca ctc ttc ctg cgc cgt cac ctg ggc tcc ggc tac tac | 2688 |
| Cys Gly Ser Pro Leu Phe Leu Arg Arg His Leu Gly Ser Gly Tyr Tyr |      |
| 885 890 895                                                     |      |
| ctg acg ctg gtg aag gcc cgc ctg ccc ctg acc acc aat gag aag gct | 2736 |
| Leu Thr Leu Val Lys Ala Arg Leu Pro Leu Thr Thr Asn Glu Lys Ala |      |
| 900 905 910                                                     |      |
| gac act gac atg gag ggc agt gtg gac acc agg cag gaa aag aag aat | 2784 |
| Asp Thr Asp Met Glu Gly Ser Val Asp Thr Arg Gln Glu Lys Lys Asn |      |
| 915 920 925                                                     |      |
| ggc agc cag ggc agc aga gtc ggc act cct cag ctg ctg gcc ctg gta | 2832 |
| Gly Ser Gln Gly Ser Arg Val Gly Thr Pro Gln Leu Leu Ala Leu Val |      |
| 930 935 940                                                     |      |
| cag cac tgg gtg ccc ggg gca cgg ctg gtg gag gag ctg cca cac gag | 2880 |
| Gln His Trp Val Pro Gly Ala Arg Leu Val Glu Glu Leu Pro His Glu |      |
| 945 950 955 960                                                 |      |
| ctg gtg ctg gtg ctg ccc tac acg ggt gcc cat gac ggc agc ttc gcc | 2928 |
| Leu Val Leu Val Leu Pro Tyr Thr Gly Ala His Asp Gly Ser Phe Ala |      |
| 965 970 975                                                     |      |
| aca ctc ttc cga gag cta gac acg cgg ctg gcg gag ctg agg ctc act | 2976 |
| Thr Leu Phe Arg Glu Leu Asp Thr Arg Leu Ala Glu Leu Arg Leu Thr |      |
| 980 985 990                                                     |      |
| ggc tac ggg atc tcc gac acc agc ctc gag gag atc ttc ctg aag gtg | 3024 |
| Gly Tyr Gly Ile Ser Asp Thr Ser Leu Glu Glu Ile Phe Leu Lys Val |      |
| 995 1000 1005                                                   |      |
| gtg gag gag tgt gct gcg gac aca gat atg gag gat ggc agc tgc ggg | 3072 |
| Val Glu Glu Cys Ala Ala Asp Thr Asp Met Glu Asp Gly Ser Cys Gly |      |
| 1010 1015 1020                                                  |      |
| cag cac cta tgc aca ggc att gct ggc cta gac gta acc ctg cgg ctc | 3120 |
| Gln His Leu Cys Thr Gly Ile Ala Gly Leu Asp Val Thr Leu Arg Leu |      |
| 1025 1030 1035 1040                                             |      |
| aag atg ccg cca cag gag aca gcg ctg gag aac ggg gaa cca gct ggg | 3168 |
| Lys Met Pro Pro Gln Glu Thr Ala Leu Glu Asn Gly Glu Pro Ala Gly |      |
| 1045 1050 1055                                                  |      |
| tca gcc cca gag act gac cag ggc tct ggg cca gac gcc gtg ggc cgg | 3216 |
| Ser Ala Pro Glu Thr Asp Gln Gly Ser Gly Pro Asp Ala Val Gly Arg |      |
| 1060 1065 1070                                                  |      |
| gta cag ggc tgg gca ctg acc cgc cag cag ctc cag gcc ctg ctt ctc | 3264 |
| Val Gln Gly Trp Ala Leu Thr Arg Gln Gln Leu Gln Ala Leu Leu Leu |      |
| 1075 1080 1085                                                  |      |
| aag cgc ttt ctg ctt gcc cgc cgc agc cgc cgc ggc ctg ttc gcc cag | 3312 |
| Lys Arg Phe Leu Leu Ala Arg Arg Ser Arg Arg Gly Leu Phe Ala Gln |      |
| 1090 1095 1100                                                  |      |
| atc gtg ctg cct gcc ctc ttt gtg ggc ctg gcc ctc gtg ttc agc ctc | 3360 |
| Ile Val Leu Pro Ala Leu Phe Val Gly Leu Ala Leu Val Phe Ser Leu |      |

| 1105                                                            | 1110 | 1115 | 1120 |      |
|-----------------------------------------------------------------|------|------|------|------|
| atc gtg cct cct ttc ggg cac tac ccg gct ctg cgg ctc agt ccc acc |      |      |      | 3408 |
| Ile Val Pro Pro Phe Gly His Tyr Pro Ala Leu Arg Leu Ser Pro Thr |      |      |      |      |
| 1125                                                            |      | 1130 | 1135 |      |
| atg tac ggt gct cag gtg tcc ttc ttc agt gag gac gcc cca ggg gac |      |      |      | 3456 |
| Met Tyr Gly Ala Gln Val Ser Phe Phe Ser Glu Asp Ala Pro Gly Asp |      |      |      |      |
| 1140                                                            |      | 1145 | 1150 |      |
| cct gga cgt gcc cgg ctg ctc gag gcg ctg ctg cag gag gca gga ctg |      |      |      | 3504 |
| Pro Gly Arg Ala Arg Leu Leu Glu Ala Leu Leu Gln Glu Ala Gly Leu |      |      |      |      |
| 1155                                                            |      | 1160 | 1165 |      |
| gag gag ccc cca gtg cag cat agc tcc cac agg ttc tcg gca cca gaa |      |      |      | 3552 |
| Glu Glu Pro Pro Val Gln His Ser Ser His Arg Phe Ser Ala Pro Glu |      |      |      |      |
| 1170                                                            |      | 1175 | 1180 |      |
| gtt cct gct gaa gtg gcc aag gtc ttg gcc agt ggc aac tgg acc cca |      |      |      | 3600 |
| Val Pro Ala Glu Val Ala Lys Val Leu Ala Ser Gly Asn Trp Thr Pro |      |      |      |      |
| 1185                                                            |      | 1190 | 1195 | 1200 |
| gag tct cca tcc cca gcc tgc cag tgt agc cag ccc ggt gcc cgg cgc |      |      |      | 3648 |
| Glu Ser Pro Ser Pro Ala Cys Gln Cys Ser Gln Pro Gly Ala Arg Arg |      |      |      |      |
| 1205                                                            |      | 1210 | 1215 |      |
| ctg ctg ccc gac tgc ccg gct gca gct ggt ggt ccc cct ccg ccc cag |      |      |      | 3696 |
| Leu Leu Pro Asp Cys Pro Ala Ala Ala Gly Gly Pro Pro Pro Gln     |      |      |      |      |
| 1220                                                            |      | 1225 | 1230 |      |
| gca gtg acc ggc tct ggg gaa gtg gtt cag aac ctg aca ggc cgg aac |      |      |      | 3744 |
| Ala Val Thr Gly Ser Gly Glu Val Val Gln Asn Leu Thr Gly Arg Asn |      |      |      |      |
| 1235                                                            |      | 1240 | 1245 |      |
| ctg tct gac ttc ctg gtc aag acc tac ccg cgc ctg gtg cgc cag ggc |      |      |      | 3792 |
| Leu Ser Asp Phe Leu Val Lys Thr Tyr Pro Arg Leu Val Arg Gln Gly |      |      |      |      |
| 1250                                                            |      | 1255 | 1260 |      |
| ctg aag act aag aag tgg gtg aat gag gtc agg tac gga ggc ttc tcg |      |      |      | 3840 |
| Leu Lys Thr Lys Lys Trp Val Asn Glu Val Arg Tyr Gly Gly Phe Ser |      |      |      |      |
| 1265                                                            |      | 1270 | 1275 | 1280 |
| ctg ggg ggc cga gac cca ggc ctg ccc tcg ggc caa gag ttg ggc cgc |      |      |      | 3888 |
| Leu Gly Gly Arg Asp Pro Gly Leu Pro Ser Gly Gln Glu Leu Gly Arg |      |      |      |      |
| 1285                                                            |      | 1290 | 1295 |      |
| tca gtg gag gag ttg tgg gcg ctg ctg agt ccc ctg cct ggc ggg gcc |      |      |      | 3936 |
| Ser Val Glu Glu Leu Trp Ala Leu Leu Ser Pro Leu Pro Gly Gly Ala |      |      |      |      |
| 1300                                                            |      | 1305 | 1310 |      |
| ctc gac cgt gtc ctg aaa aac ctc aca gcc tgg gct cac agc ctg gat |      |      |      | 3984 |
| Leu Asp Arg Val Leu Lys Asn Leu Thr Ala Trp Ala His Ser Leu Asp |      |      |      |      |
| 1315                                                            |      | 1320 | 1325 |      |
| gct cag gac agt ctc aag atc tgg ttc aac aac aaa ggc tgg cac tcc |      |      |      | 4032 |
| Ala Gln Asp Ser Leu Lys Ile Trp Phe Asn Asn Lys Gly Trp His Ser |      |      |      |      |
| 1330                                                            |      | 1335 | 1340 |      |

```

atg gtg gcc ttt gtc aac cga gcc agc aac gca atc ctc cgt gct cac 4080
Met Val Ala Phe Val Asn Arg Ala Ser Asn Ala Ile Leu Arg Ala His
1345          1350          1355          1360

ctg ccc cca ggc ccg gcc cgc cac gcc cac agc atc acc aca ctc aac 4128
Leu Pro Pro Gly Pro Ala Arg His Ala His Ser Ile Thr Thr Leu Asn
          1365          1370          1375

cac ccc ttg aac ctc acc aag gag cag ctg tct gag gct gca ctg atg 4176
His Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ala Ala Leu Met
          1380          1385          1390

gcc tcc tcg gtg gac gtc ctc gtc tcc atc tgt gtg gtc ttt gcc atg 4224
Ala Ser Ser Ser Val Asp Val Leu Val Ser Ile Cys Val Val Phe Ala Met
          1395          1400          1405

tcc ttt gtc ccg gcc agc ttc act ctt gtc ctc att gag gag cga gtc 4272
Ser Phe Val Pro Ala Ser Phe Thr Leu Val Leu Ile Glu Glu Arg Val
          1410          1415          1420

acc cga gcc aag cac ctg cag ctc atg ggg ggc ctg tcc ccc acc ctc 4320
Thr Arg Ala Lys His Leu Gln Leu Met Gly Gly Leu Ser Pro Thr Leu
1425          1430          1435          1440

tac tgg ctt ggc aac ttt ctc tgg gac atg tgt aac tac ttg gtg cca 4368
Tyr Trp Leu Gly Asn Phe Leu Trp Asp Met Cys Asn Tyr Leu Val Pro
          1445          1450          1455

gca tgc atc gtg gtg ctc atc ttt ctg gcc ttc cag cag agg gca tat 4416
Ala Cys Ile Val Val Leu Ile Phe Leu Ala Phe Gln Gln Arg Ala Tyr
          1460          1465          1470

gtg gcc cct gcc aac ctg cct gct ctc ctg ctg ttg cta cta ctg tat 4464
Val Ala Pro Ala Asn Leu Pro Ala Leu Leu Leu Leu Leu Leu Tyr
          1475          1480          1485

ggg aga cag gca gtt cca gtc acc cct gcg ctg gga ggt ggt cgg caa 4512
Gly Arg Gln Ala Val Pro Val Thr Pro Ala Leu Gly Gly Gly Arg Gln
1490          1495          1500

gaa cct ctt ggc cat ggt gat aca ggg gcc cct ctt cct tct ctt cac 4560
Glu Pro Leu Gly His Gly Asp Thr Gly Ala Pro Leu Pro Ser Leu His
1505          1510          1515          1520

act act gct gca gca ccg aag cca act cct gcc aca gcc cag ggt gag 4608
Thr Thr Ala Ala Ala Pro Lys Pro Thr Pro Ala Thr Ala Gln Gly Glu
          1525          1530          1535

gtc tct gcc act cct ggg aga gga gga cga gga tgt agc ccg tga 4653
Val Ser Ala Thr Pro Gly Arg Gly Gly Arg Gly Cys Ser Pro
          1540          1545          1550

```

&lt;210&gt; 8

&lt;211&gt; 1550

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 8

```

Met Val Cys Leu Gly Thr Gly Gln Ser Ala Gly Pro Leu Val Ser Val
 1           5           10           15

Gln Asn His Cys Pro Pro Cys Gly Leu Ser Pro Gln Glu Ser Leu Gly
          20           25           30

Leu Ala Leu Gly Gln Ala Gln Glu Pro Leu His Ser Leu Leu Glu Ala
      35           40           45

Ala Gly Asp Leu Ala Gln Glu Leu Leu Ala Leu Arg Ser Leu Val Glu
      50           55           60

Leu Arg Ala Leu Leu Gln Arg Pro Arg Gly Thr Ser Gly Pro Leu Glu
      65           70           75           80

Leu Leu Ser Glu Ala Leu Cys Ser Val Arg Gly Pro Ser Ser Thr Val
          85           90           95

Gly Pro Ser Leu Asn Trp Tyr Glu Ala Ser Asp Leu Met Glu Leu Val
          100           105           110

Gly Gln Glu Pro Glu Ser Ala Leu Pro Asp Ser Ser Leu Ser Pro Ala
      115           120           125

Cys Ser Glu Leu Ile Gly Ala Leu Asp Ser His Pro Leu Ser Arg Leu
      130           135           140

Leu Trp Arg Arg Leu Lys Pro Leu Ile Leu Gly Lys Leu Leu Phe Ala
      145           150           155           160

Pro Asp Thr Pro Phe Thr Arg Lys Leu Met Ala Gln Val Asn Arg Thr
          165           170           175

Phe Glu Glu Leu Thr Leu Leu Arg Asp Val Arg Glu Val Trp Glu Met
          180           185           190

Leu Gly Pro Arg Ile Phe Thr Phe Met Asn Asp Ser Ser Asn Val Ala
      195           200           205

Met Leu Gln Arg Leu Leu Gln Met Gln Asp Glu Gly Arg Arg Gln Pro
      210           215           220

Arg Pro Gly Gly Arg Asp His Met Glu Ala Leu Arg Ser Phe Leu Asp
      225           230           235           240

Pro Gly Ser Gly Gly Tyr Ser Trp Gln Asp Ala His Ala Asp Val Gly
          245           250           255

His Leu Val Gly Thr Leu Gly Arg Val Thr Glu Cys Leu Ser Leu Asp
          260           265           270

Lys Leu Glu Ala Ala Pro Ser Glu Ala Ala Leu Val Ser Arg Ala Leu
      275           280           285

Gln Leu Leu Ala Glu His Arg Phe Trp Ala Gly Val Val Phe Leu Gly
      290           295           300

```

Pro Glu Asp Ser Ser Asp Pro Thr Glu His Pro Thr Pro Asp Leu Gly  
 305 310 315 320  
 Pro Gly His Val Arg Ile Lys Ile Arg Met Asp Ile Asp Val Val Thr  
 325 330 335  
 Arg Thr Asn Lys Ile Arg Asp Arg Phe Trp Asp Pro Gly Pro Ala Ala  
 340 345 350  
 Asp Pro Leu Thr Asp Leu Arg Tyr Val Trp Gly Gly Phe Val Tyr Leu  
 355 360 365  
 Gln Asp Leu Val Glu Arg Ala Ala Val Arg Val Leu Ser Gly Ala Asn  
 370 375 380  
 Pro Arg Ala Gly Leu Tyr Leu Gln Gln Met Pro Tyr Pro Cys Tyr Val  
 385 390 395 400  
 Asp Asp Val Phe Leu Arg Val Leu Ser Arg Ser Leu Pro Leu Phe Leu  
 405 410 415  
 Thr Leu Ala Trp Ile Tyr Ser Val Thr Leu Thr Val Lys Ala Val Val  
 420 425 430  
 Arg Glu Lys Glu Thr Arg Leu Arg Asp Thr Met Arg Ala Met Gly Leu  
 435 440 445  
 Ser Arg Ala Val Leu Trp Leu Gly Trp Phe Leu Ser Cys Leu Gly Pro  
 450 455 460  
 Phe Leu Leu Ser Ala Ala Leu Leu Val Leu Val Leu Lys Leu Gly Asp  
 465 470 475 480  
 Ile Leu Pro Tyr Ser His Pro Gly Val Val Phe Leu Phe Leu Ala Ala  
 485 490 495  
 Phe Ala Val Ala Thr Val Thr Gln Ser Phe Leu Leu Ser Ala Phe Phe  
 500 505 510  
 Ser Arg Ala Asn Leu Ala Ala Ala Cys Gly Gly Leu Ala Tyr Phe Ser  
 515 520 525  
 Leu Tyr Leu Pro Tyr Val Leu Cys Val Ala Trp Arg Asp Arg Leu Pro  
 530 535 540  
 Ala Gly Gly Arg Val Ala Ala Ser Leu Leu Ser Pro Val Ala Phe Gly  
 545 550 555 560  
 Phe Gly Cys Glu Ser Leu Ala Leu Leu Glu Glu Gln Gly Glu Gly Ala  
 565 570 575  
 Gln Trp His Asn Val Gly Thr Arg Pro Thr Ala Asp Val Phe Ser Leu  
 580 585 590  
 Ala Gln Val Ser Gly Leu Leu Leu Leu Asp Ala Ala Leu Tyr Gly Leu  
 595 600 605  
 Ala Thr Trp Tyr Leu Glu Ala Val Cys Pro Gly Gln Tyr Gly Ile Pro

| 610        |     |     |            |            | 615        |            |            |            |            | 620        |            |            |            |     |            |
|------------|-----|-----|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-----|------------|
| Glu<br>625 | Pro | Trp | Asn        | Phe        | Pro<br>630 | Phe        | Arg        | Arg        | Ser        | Tyr<br>635 | Trp        | Cys        | Gly        | Pro | Arg<br>640 |
| Pro        | Pro | Lys | Ser        | Pro<br>645 | Ala        | Pro        | Cys        | Pro        | Thr<br>650 | Pro        | Leu        | Asp        | Pro        | Lys | Val<br>655 |
| Leu        | Val | Glu | Glu        | Ala<br>660 | Pro        | Pro        | Gly        | Leu        | Ser<br>665 | Pro        | Gly        | Val        | Ser        | Val | Arg<br>670 |
| Ser        | Leu | Glu | Lys        | Arg        | Phe        | Pro        | Gly<br>680 | Ser        | Pro        | Gln        | Pro        | Ala<br>685 | Leu        | Arg | Gly<br>690 |
| Leu<br>690 | Ser | Leu | Asp        | Phe        | Tyr        | Gln<br>695 | Gly        | His        | Ile        | Thr        | Ala<br>700 | Phe        | Leu        | Gly | His<br>705 |
| Asn<br>705 | Gly | Ala | Gly        | Lys        | Thr<br>710 | Thr        | Thr        | Leu        | Ser        | Ile<br>715 | Leu        | Ser        | Gly        | Leu | Phe<br>720 |
| Pro        | Pro | Ser | Gly        | Gly<br>725 | Ser        | Ala        | Phe        | Ile        | Leu        | Gly<br>730 | His        | Asp        | Val        | Arg | Ser<br>735 |
| Ser        | Met | Ala | Ala        | Ile<br>740 | Arg        | Pro        | His        | Leu        | Gly<br>745 | Val        | Cys        | Pro        | Gln        | Tyr | Asn<br>750 |
| Val        | Leu | Phe | Asp        | Met<br>755 | Leu        | Thr        | Val        | Asp        | Glu        | His<br>760 | Val        | Trp        | Phe        | Tyr | Gly<br>765 |
| Arg<br>770 | Leu | Lys | Gly        | Leu        | Ser        | Ala<br>775 | Ala        | Val        | Val        | Gly        | Pro<br>780 | Glu        | Gln        | Asp | Arg<br>785 |
| Leu<br>785 | Leu | Gln | Asp        | Val        | Gly<br>790 | Leu        | Val        | Ser        | Lys        | Gln<br>795 | Ser        | Val        | Gln        | Thr | Arg<br>800 |
| His        | Leu | Ser | Gly        | Gly<br>805 | Met        | Gln        | Arg        | Lys        | Leu<br>810 | Ser        | Val        | Ala        | Ile        | Ala | Phe<br>815 |
| Val        | Gly | Gly | Ser        | Gln<br>820 | Val        | Val        | Ile        | Leu        | Asp<br>825 | Glu        | Pro        | Thr        | Ala        | Gly | Val<br>830 |
| Asp        | Pro | Ala | Ser        | Arg        | Arg        | Gly        | Ile<br>840 | Trp        | Glu        | Leu        | Leu<br>845 | Leu        | Lys        | Tyr | Arg<br>850 |
| Glu<br>850 | Gly | Arg | Thr        | Leu        | Ile        | Leu<br>855 | Ser        | Thr        | His        | His        | Leu<br>860 | Asp        | Glu        | Ala | Glu<br>865 |
| Leu<br>865 | Leu | Gly | Asp        | Arg        | Val<br>870 | Ala        | Val        | Val        | Ala        | Gly<br>875 | Gly        | Arg        | Leu        | Cys | Cys<br>880 |
| Cys        | Gly | Ser | Pro        | Leu<br>885 | Phe        | Leu        | Arg        | Arg        | His<br>890 | Leu        | Gly        | Ser        | Gly        | Tyr | Tyr<br>895 |
| Leu        | Thr | Leu | Val<br>900 | Lys        | Ala        | Arg        | Leu        | Pro<br>905 | Leu        | Thr        | Thr        | Asn        | Glu<br>910 | Lys | Ala<br>915 |
| Asp        | Thr | Asp | Met        | Glu        | Gly        | Ser        | Val<br>920 | Asp        | Thr        | Arg        | Gln        | Glu<br>925 | Lys        | Lys | Asn<br>930 |

Gly Ser Gln Gly Ser Arg Val Gly Thr Pro Gln Leu Leu Ala Leu Val  
 930 935 940  
 Gln His Trp Val Pro Gly Ala Arg Leu Val Glu Glu Leu Pro His Glu  
 945 950 955 960  
 Leu Val Leu Val Leu Pro Tyr Thr Gly Ala His Asp Gly Ser Phe Ala  
 965 970 975  
 Thr Leu Phe Arg Glu Leu Asp Thr Arg Leu Ala Glu Leu Arg Leu Thr  
 980 985 990  
 Gly Tyr Gly Ile Ser Asp Thr Ser Leu Glu Glu Ile Phe Leu Lys Val  
 995 1000 1005  
 Val Glu Glu Cys Ala Ala Asp Thr Asp Met Glu Asp Gly Ser Cys Gly  
 1010 1015 1020  
 Gln His Leu Cys Thr Gly Ile Ala Gly Leu Asp Val Thr Leu Arg Leu  
 1025 1030 1035 1040  
 Lys Met Pro Pro Gln Glu Thr Ala Leu Glu Asn Gly Glu Pro Ala Gly  
 1045 1050 1055  
 Ser Ala Pro Glu Thr Asp Gln Gly Ser Gly Pro Asp Ala Val Gly Arg  
 1060 1065 1070  
 Val Gln Gly Trp Ala Leu Thr Arg Gln Gln Leu Gln Ala Leu Leu Leu  
 1075 1080 1085  
 Lys Arg Phe Leu Leu Ala Arg Arg Ser Arg Arg Gly Leu Phe Ala Gln  
 1090 1095 1100  
 Ile Val Leu Pro Ala Leu Phe Val Gly Leu Ala Leu Val Phe Ser Leu  
 1105 1110 1115 1120  
 Ile Val Pro Pro Phe Gly His Tyr Pro Ala Leu Arg Leu Ser Pro Thr  
 1125 1130 1135  
 Met Tyr Gly Ala Gln Val Ser Phe Phe Ser Glu Asp Ala Pro Gly Asp  
 1140 1145 1150  
 Pro Gly Arg Ala Arg Leu Leu Glu Ala Leu Leu Gln Glu Ala Gly Leu  
 1155 1160 1165  
 Glu Glu Pro Pro Val Gln His Ser Ser His Arg Phe Ser Ala Pro Glu  
 1170 1175 1180  
 Val Pro Ala Glu Val Ala Lys Val Leu Ala Ser Gly Asn Trp Thr Pro  
 1185 1190 1195 1200  
 Glu Ser Pro Ser Pro Ala Cys Gln Cys Ser Gln Pro Gly Ala Arg Arg  
 1205 1210 1215  
 Leu Leu Pro Asp Cys Pro Ala Ala Ala Gly Gly Pro Pro Pro Pro Gln  
 1220 1225 1230

Ala Val Thr Gly Ser Gly Glu Val Val Gln Asn Leu Thr Gly Arg Asn  
 1235 1240 1245  
 Leu Ser Asp Phe Leu Val Lys Thr Tyr Pro Arg Leu Val Arg Gln Gly  
 1250 1255 1260  
 Leu Lys Thr Lys Lys Trp Val Asn Glu Val Arg Tyr Gly Gly Phe Ser  
 1265 1270 1275 1280  
 Leu Gly Gly Arg Asp Pro Gly Leu Pro Ser Gly Gln Glu Leu Gly Arg  
 1285 1290 1295  
 Ser Val Glu Glu Leu Trp Ala Leu Leu Ser Pro Leu Pro Gly Gly Ala  
 1300 1305 1310  
 Leu Asp Arg Val Leu Lys Asn Leu Thr Ala Trp Ala His Ser Leu Asp  
 1315 1320 1325  
 Ala Gln Asp Ser Leu Lys Ile Trp Phe Asn Asn Lys Gly Trp His Ser  
 1330 1335 1340  
 Met Val Ala Phe Val Asn Arg Ala Ser Asn Ala Ile Leu Arg Ala His  
 1345 1350 1355 1360  
 Leu Pro Pro Gly Pro Ala Arg His Ala His Ser Ile Thr Thr Leu Asn  
 1365 1370 1375  
 His Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu Ala Ala Leu Met  
 1380 1385 1390  
 Ala Ser Ser Val Asp Val Leu Val Ser Ile Cys Val Val Phe Ala Met  
 1395 1400 1405  
 Ser Phe Val Pro Ala Ser Phe Thr Leu Val Leu Ile Glu Glu Arg Val  
 1410 1415 1420  
 Thr Arg Ala Lys His Leu Gln Leu Met Gly Gly Leu Ser Pro Thr Leu  
 1425 1430 1435 1440  
 Tyr Trp Leu Gly Asn Phe Leu Trp Asp Met Cys Asn Tyr Leu Val Pro  
 1445 1450 1455  
 Ala Cys Ile Val Val Leu Ile Phe Leu Ala Phe Gln Gln Arg Ala Tyr  
 1460 1465 1470  
 Val Ala Pro Ala Asn Leu Pro Ala Leu Leu Leu Leu Leu Leu Tyr  
 1475 1480 1485  
 Gly Arg Gln Ala Val Pro Val Thr Pro Ala Leu Gly Gly Gly Arg Gln  
 1490 1495 1500  
 Glu Pro Leu Gly His Gly Asp Thr Gly Ala Pro Leu Pro Ser Leu His  
 1505 1510 1515 1520  
 Thr Thr Ala Ala Ala Pro Lys Pro Thr Pro Ala Thr Ala Gln Gly Glu  
 1525 1530 1535  
 Val Ser Ala Thr Pro Gly Arg Gly Gly Arg Gly Cys Ser Pro

1540

1545

1550

&lt;210&gt; 9

&lt;211&gt; 2201

&lt;212&gt; PRT

&lt;213&gt; Mus musculus

&lt;220&gt;

&lt;221&gt; UNSURE

&lt;222&gt; (115)

&lt;223&gt; amino acid at this position is unknown

&lt;400&gt; 9

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Pro | Ser | Ala | Gly | Thr | Leu | Pro | Trp | Val | Gln | Gly | Ile | Ile | Cys | Asn |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Asn | Asn | Pro | Cys | Phe | Arg | Tyr | Pro | Thr | Pro | Gly | Glu | Ala | Pro | Gly |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Val | Gly | Asn | Phe | Asn | Lys | Ser | Ile | Val | Ser | Arg | Leu | Phe | Ser | Asp |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Gln | Arg | Leu | Leu | Leu | Tyr | Ser | Gln | Arg | Asp | Thr | Ser | Ile | Lys | Asp |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | His | Lys | Val | Leu | Arg | Met | Leu | Arg | Gln | Ile | Lys | His | Pro | Asn | Ser |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Leu | Lys | Leu | Gln | Asp | Phe | Leu | Val | Asp | Asn | Glu | Thr | Phe | Ser | Gly |
|     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Phe | Leu | Gln | His | Asn | Leu | Ser | Leu | Pro | Arg | Ser | Thr | Val | Asp | Ser | Leu |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Gln | Xaa | Asn | Val | Gly | Leu | Gln | Lys | Val | Phe | Leu | Gln | Gly | Tyr | Gln |
|     | 115 |     |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | His | Leu | Ala | Ser | Leu | Cys | Asn | Gly | Ser | Lys | Leu | Glu | Glu | Ile | Ile |
|     | 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gln | Leu | Gly | Asp | Ala | Glu | Val | Ser | Ala | Leu | Cys | Gly | Leu | Pro | Arg | Lys |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lys | Leu | Asp | Ala | Ala | Glu | Arg | Val | Leu | Arg | Tyr | Asn | Met | Asp | Ile | Leu |
|     |     |     | 165 |     |     |     |     |     | 170 |     |     |     |     | 175 |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lys | Pro | Val | Val | Thr | Lys | Leu | Asn | Ser | Thr | Ser | His | Leu | Pro | Thr | Gln |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     | 190 |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| His | Leu | Ala | Glu | Ala | Thr | Thr | Val | Leu | Leu | Asp | Ser | Leu | Gly | Gly | Leu |
|     | 195 |     |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Gln | Glu | Leu | Phe | Ser | Thr | Lys | Ser | Trp | Ser | Asp | Met | Arg | Gln | Glu |
|     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Met | Phe | Leu | Thr | Asn | Val | Asn | Ser | Ser | Ser | Ser | Ser | Thr | Gln | Ile |
| 225 |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |

Tyr Gln Ala Val Ser Arg Ile Val Cys Gly His Pro Glu Gly Gly Gly  
 245 250 255  
 Leu Lys Ile Lys Ser Leu Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala  
 260 265 270  
 Leu Phe Gly Gly Asn Asn Thr Glu Glu Asp Val Asp Thr Phe Tyr Asp  
 275 280 285  
 Asn Ser Thr Thr Pro Tyr Cys Asn Asp Leu Met Lys Asn Leu Glu Ser  
 290 295 300  
 Ser Pro Leu Ser Arg Ile Ile Trp Lys Ala Leu Lys Pro Leu Leu Val  
 305 310 315 320  
 Gly Lys Ile Leu Tyr Thr Pro Asp Thr Pro Ala Thr Arg Gln Val Met  
 325 330 335  
 Ala Glu Val Asn Lys Thr Phe Gln Glu Leu Ala Val Phe His Asp Leu  
 340 345 350  
 Glu Gly Met Trp Glu Glu Leu Ser Pro Gln Ile Trp Thr Phe Met Glu  
 355 360 365  
 Asn Ser Gln Glu Met Asp Leu Val Arg Thr Leu Leu Asp Ser Arg Gly  
 370 375 380  
 Asn Asp Gln Phe Trp Glu Gln Lys Leu Asp Gly Leu Asp Trp Thr Ala  
 385 390 395 400  
 Gln Asp Ile Met Ala Phe Leu Ala Lys Asn Pro Glu Asp Val Gln Ser  
 405 410 415  
 Pro Asn Gly Ser Val Tyr Thr Trp Arg Glu Ala Phe Asn Glu Thr Asn  
 420 425 430  
 Gln Ala Ile Gln Thr Ile Ser Arg Phe Met Glu Cys Val Asn Leu Asn  
 435 440 445  
 Lys Leu Glu Pro Ile Pro Thr Glu Val Arg Leu Ile Asn Lys Ser Met  
 450 455 460  
 Glu Leu Leu Asp Glu Arg Lys Phe Trp Ala Gly Ile Val Phe Thr Gly  
 465 470 475 480  
 Ile Thr Pro Asp Ser Val Glu Leu Pro His His Val Lys Tyr Lys Ile  
 485 490 495  
 Arg Met Asp Ile Asp Asn Val Glu Arg Thr Asn Lys Ile Lys Asp Gly  
 500 505 510  
 Tyr Trp Asp Pro Gly Pro Arg Ala Asp Pro Phe Glu Asp Met Arg Tyr  
 515 520 525  
 Val Trp Gly Gly Phe Ala Tyr Leu Gln Asp Val Val Glu Gln Ala Ile  
 530 535 540

```

Ile Arg Val Leu Thr Gly Ser Glu Lys Lys Thr Gly Val Tyr Val Gln
545                550                555                560

Gln Met Pro Tyr Pro Cys Tyr Val Asp Asp Ile Phe Leu Arg Val Met
565                570

Ser Arg Ser Met Pro Leu Phe Met Thr Leu Ala Trp Ile Tyr Ser Val
580                585                590

Ala Val Ile Ile Lys Ser Ile Val Tyr Glu Lys Glu Ala Arg Leu Lys
595                600                605

Glu Thr Met Arg Ile Met Gly Leu Asp Asn Gly Ile Leu Trp Phe Ser
610                615                620

Trp Phe Val Ser Ser Leu Ile Pro Leu Leu Val Ser Ala Gly Leu Leu
625                630                635

Val Val Ile Leu Lys Leu Gly Asn Leu Leu Pro Tyr Ser Asp Pro Ser
645                650                655

Val Val Phe Val Phe Leu Ser Val Phe Ala Met Val Thr Ile Leu Gln
660                665                670

Cys Phe Leu Ile Ser Thr Leu Phe Ser Arg Ala Asn Leu Ala Ala Ala
675                680                685

Cys Gly Gly Ile Ile Tyr Phe Thr Leu Tyr Leu Pro Tyr Val Leu Cys
690                695                700

Val Ala Trp Gln Asp Tyr Val Gly Phe Ser Ile Lys Ile Phe Ala Ser
705                710                715

Leu Leu Ser Pro Val Ala Phe Gly Phe Gly Cys Glu Tyr Phe Ala Leu
725                730                735

Phe Glu Glu Gln Gly Ile Gly Val Gln Trp Asp Asn Leu Phe Glu Ser
740                745                750

Pro Val Glu Glu Asp Gly Phe Asn Leu Thr Thr Ala Val Ser Met Met
755                760                765

Leu Phe Asp Thr Phe Leu Tyr Gly Val Met Thr Trp Tyr Ile Glu Ala
770                775                780

Val Phe Pro Gly Gln Tyr Gly Ile Pro Arg Pro Trp Tyr Phe Pro Cys
785                790                795

Thr Lys Ser Tyr Trp Phe Gly Glu Glu Ile Asp Glu Lys Ser His Pro
805                810                815

Gly Ser Ser Gln Lys Gly Val Ser Glu Ile Cys Met Glu Glu Glu Pro
820                825                830

Thr His Leu Arg Leu Gly Val Ser Ile Gln Asn Leu Val Lys Val Tyr
835                840                845

Arg Asp Gly Met Lys Val Ala Val Asp Gly Leu Ala Leu Asn Phe Tyr

```

| 850                                                                                    | 855 | 860 |
|----------------------------------------------------------------------------------------|-----|-----|
| Glu Gly Gln Ile Thr Ser Phe Leu Gly His Asn Gly Ala Gly Lys Thr<br>865 870 875 880     |     |     |
| Thr Thr Met Ser Ile Leu Thr Gly Leu Phe Pro Pro Thr Ser Gly Thr<br>885 890 895         |     |     |
| Ala Tyr Ile Leu Gly Lys Asp Ile Arg Ser Glu Met Ser Ser Ile Arg<br>900 905 910         |     |     |
| Gln Asn Leu Gly Val Cys Pro Gln His Asn Val Leu Phe Asp Met Leu<br>915 920 925         |     |     |
| Thr Val Glu Glu His Ile Trp Phe Tyr Ala Arg Leu Lys Gly Leu Ser<br>930 935 940         |     |     |
| Glu Lys His Val Lys Ala Glu Met Glu Gln Met Ala Leu Asp Val Gly<br>945 950 955 960     |     |     |
| Leu Pro Pro Ser Lys Leu Lys Ser Lys Thr Ser Gln Leu Ser Gly Gly<br>965 970 975         |     |     |
| Met Gln Arg Lys Leu Ser Val Ala Leu Ala Phe Val Gly Gly Ser Lys<br>980 985 990         |     |     |
| Val Val Ile Leu Asp Glu Pro Thr Ala Gly Val Asp Pro Tyr Ser Arg<br>995 1000 1005       |     |     |
| Arg Gly Ile Trp Glu Leu Leu Leu Lys Tyr Arg Gln Gly Arg Thr Ile<br>1010 1015 1020      |     |     |
| Ile Leu Ser Thr His His Met Asp Glu Ala Asp Ile Leu Gly Asp Arg<br>1025 1030 1035 1040 |     |     |
| Ile Ala Ile Ile Ser His Gly Lys Leu Cys Cys Val Gly Ser Ser Leu<br>1045 1050 1055      |     |     |
| Phe Leu Lys Asn Gln Leu Gly Thr Gly Tyr Tyr Leu Thr Leu Val Lys<br>1060 1065 1070      |     |     |
| Lys Asp Val Glu Ser Ser Leu Ser Ser Cys Arg Asn Ser Ser Ser Thr<br>1075 1080 1085      |     |     |
| Val Ser Cys Leu Lys Lys Glu Asp Ser Val Ser Gln Ser Ser Ser Asp<br>1090 1095 1100      |     |     |
| Ala Gly Leu Gly Ser Asp His Glu Ser Asp Thr Leu Thr Ile Asp Val<br>1105 1110 1115 1120 |     |     |
| Ser Ala Ile Ser Asn Leu Ile Arg Lys His Val Ser Glu Ala Arg Leu<br>1125 1130 1135      |     |     |
| Val Glu Asp Ile Gly His Glu Leu Thr Tyr Val Leu Pro Tyr Glu Ala<br>1140 1145 1150      |     |     |
| Ala Lys Glu Gly Ala Phe Val Glu Leu Phe His Glu Ile Asp Asp Arg<br>1155 1160 1165      |     |     |

Leu Ser Asp Leu Gly Ile Ser Ser Tyr Gly Ile Ser Glu Thr Thr Leu  
 1170 1175 1180  
 Glu Glu Ile Phe Leu Lys Val Ala Glu Glu Ser Gly Val Asp Ala Glu  
 1185 1190 1195 1200  
 Thr Ser Asp Gly Thr Leu Pro Ala Arg Arg Asn Arg Arg Ala Phe Gly  
 1205 1210 1215  
 Asp Lys Gln Ser Cys Leu His Pro Phe Thr Glu Asp Asp Ala Val Asp  
 1220 1225 1230  
 Pro Asn Asp Ser Asp Ile Asp Pro Glu Ser Arg Glu Thr Asp Leu Leu  
 1235 1240 1245  
 Ser Gly Met Asp Gly Lys Gly Ser Tyr Gln Leu Lys Gly Trp Lys Leu  
 1250 1255 1260  
 Thr Gln Gln Gln Phe Val Ala Leu Leu Trp Lys Arg Leu Leu Ile Ala  
 1265 1270 1275 1280  
 Arg Arg Ser Arg Lys Gly Phe Phe Ala Gln Ile Val Leu Pro Ala Val  
 1285 1290 1295  
 Phe Val Cys Ile Ala Leu Val Phe Ser Leu Ile Val Pro Pro Phe Gly  
 1300 1305 1310  
 Lys Tyr Pro Ser Leu Glu Leu Gln Pro Trp Met Tyr Asn Glu Gln Tyr  
 1315 1320 1325  
 Thr Phe Val Ser Asn Asp Ala Pro Glu Asp Met Gly Thr Gln Glu Leu  
 1330 1335 1340  
 Leu Asn Ala Leu Thr Lys Asp Pro Gly Phe Gly Thr Arg Cys Met Glu  
 1345 1350 1355 1360  
 Gly Asn Pro Ile Pro Asp Thr Pro Cys Leu Ala Gly Glu Glu Asp Trp  
 1365 1370 1375  
 Thr Ile Ser Pro Val Pro Gln Ser Ile Val Asp Leu Phe Gln Asn Gly  
 1380 1385 1390  
 Asn Trp Thr Met Lys Asn Pro Ser Pro Ala Cys Gln Cys Ser Ser Asp  
 1395 1400 1405  
 Lys Ile Lys Lys Met Leu Pro Val Cys Pro Pro Gly Ala Gly Gly Leu  
 1410 1415 1420  
 Pro Pro Pro Gln Arg Lys Gln Lys Thr Ala Asp Ile Leu Gln Asn Leu  
 1425 1430 1435 1440  
 Thr Gly Arg Asn Ile Ser Asp Tyr Leu Val Lys Thr Tyr Val Gln Ile  
 1445 1450 1455  
 Ile Ala Lys Ser Leu Lys Asn Lys Ile Trp Val Asn Glu Phe Arg Tyr  
 1460 1465 1470

Gly Gly Phe Ser Leu Gly Val Ser Asn Ser Gln Ala Leu Pro Pro Ser  
 1475 1480 1485  
 His Glu Val Asn Asp Ala Ile Lys Gln Met Lys Lys Leu Leu Lys Leu  
 1490 1495 1500  
 Thr Lys Asp Thr Ser Ala Asp Arg Phe Leu Ser Ser Leu Gly Arg Phe  
 1505 1510 1515 1520  
 Met Ala Gly Leu Asp Thr Lys Asn Asn Val Lys Val Trp Phe Asn Asn  
 1525 1530 1535  
 Lys Gly Trp His Ala Ile Ser Ser Phe Leu Asn Val Ile Asn Asn Ala  
 1540 1545 1550  
 Ile Leu Arg Ala Asn Leu Gln Lys Gly Glu Asn Pro Ser Gln Tyr Gly  
 1555 1560 1565  
 Ile Thr Ala Phe Asn His Pro Leu Asn Leu Thr Lys Gln Gln Leu Ser  
 1570 1575 1580  
 Glu Val Ala Leu Met Thr Thr Ser Val Asp Val Leu Val Ser Ile Cys  
 1585 1590 1595 1600  
 Val Ile Phe Ala Met Ser Phe Val Pro Ala Ser Phe Val Val Phe Leu  
 1605 1610 1615  
 Ile Gln Glu Arg Val Ser Lys Ala Lys His Leu Gln Phe Ile Ser Gly  
 1620 1625 1630  
 Val Lys Pro Val Ile Tyr Trp Leu Ser Asn Phe Val Trp Asp Met Cys  
 1635 1640 1645  
 Asn Tyr Val Val Pro Ala Thr Leu Val Ile Ile Ile Phe Ile Cys Phe  
 1650 1655 1660  
 Gln Gln Lys Ser Tyr Val Ser Ser Thr Asn Leu Pro Val Leu Ala Leu  
 1665 1670 1675 1680  
 Leu Leu Leu Leu Tyr Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala  
 1685 1690 1695  
 Ser Phe Val Phe Lys Ile Pro Ser Thr Ala Tyr Val Val Leu Thr Ser  
 1700 1705 1710  
 Val Asn Leu Phe Ile Gly Ile Asn Gly Ser Val Ala Thr Phe Val Leu  
 1715 1720 1725  
 Glu Leu Phe Thr Asn Asn Lys Leu Asn Asp Ile Asn Asp Ile Leu Lys  
 1730 1735 1740  
 Ser Val Phe Leu Ile Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile  
 1745 1750 1755 1760  
 Asp Met Val Lys Asn Gln Ala Met Ala Asp Ala Leu Glu Arg Phe Gly  
 1765 1770 1775  
 Glu Asn Arg Phe Val Ser Pro Leu Ser Trp Asp Leu Val Gly Arg Asn

| 1780                                                                                   | 1785 | 1790 |
|----------------------------------------------------------------------------------------|------|------|
| Leu Phe Ala Met Ala Val Glu Gly Val Val Phe Phe Leu Ile Thr Val<br>1795 1800 1805      |      |      |
| Leu Ile Gln Tyr Arg Phe Phe Ile Arg Pro Arg Pro Val Lys Ala Lys<br>1810 1815 1820      |      |      |
| Leu Pro Pro Leu Asn Asp Glu Asp Glu Asp Val Arg Arg Glu Arg Gln<br>1825 1830 1835 1840 |      |      |
| Arg Ile Leu Asp Gly Gly Gly Gln Asn Asp Ile Leu Glu Ile Lys Glu<br>1845 1850 1855      |      |      |
| Leu Thr Lys Ile Tyr Arg Arg Lys Arg Lys Pro Ala Val Asp Arg Ile<br>1860 1865 1870      |      |      |
| Cys Ile Gly Ile Pro Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn<br>1875 1880 1885      |      |      |
| Gly Ala Gly Lys Ser Thr Thr Phe Lys Met Leu Thr Gly Asp Thr Pro<br>1890 1895 1900      |      |      |
| Val Thr Arg Gly Asp Ala Phe Leu Asn Lys Asn Ser Ile Leu Ser Asn<br>1905 1910 1915 1920 |      |      |
| Ile His Glu Val His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala<br>1925 1930 1935      |      |      |
| Ile Thr Glu Leu Leu Thr Gly Arg Glu His Val Glu Phe Phe Ala Leu<br>1940 1945 1950      |      |      |
| Leu Arg Gly Val Pro Glu Lys Glu Val Gly Lys Phe Gly Glu Trp Ala<br>1955 1960 1965      |      |      |
| Ile Arg Lys Leu Gly Leu Val Lys Tyr Gly Glu Lys Tyr Ala Ser Asn<br>1970 1975 1980      |      |      |
| Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser Thr Ala Met Ala Leu Ile<br>1985 1990 1995 2000 |      |      |
| Gly Gly Pro Pro Val Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp<br>2005 2010 2015      |      |      |
| Pro Lys Ala Arg Arg Phe Leu Trp Asn Cys Ala Leu Ser Ile Val Lys<br>2020 2025 2030      |      |      |
| Glu Gly Arg Ser Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu<br>2035 2040 2045      |      |      |
| Ala Leu Cys Thr Arg Met Ala Ile Met Val Asn Gly Arg Phe Arg Cys<br>2050 2055 2060      |      |      |
| Leu Gly Ser Val Gln His Leu Lys Asn Arg Phe Gly Asp Gly Tyr Thr<br>2065 2070 2075 2080 |      |      |
| Ile Val Val Arg Ile Ala Gly Ser Asn Pro Asp Leu Lys Pro Val Gln<br>2085 2090 2095      |      |      |

Glu Phe Phe Gly Leu Ala Phe Pro Gly Ser Val Leu Lys Glu Lys His  
 2100 2105 2110

Arg Asn Met Leu Gln Tyr Gln Leu Pro Ser Ser Leu Ser Ser Leu Ala  
 2115 2120 2125

Arg Ile Phe Ser Ile Leu Ser Gln Ser Lys Lys Arg Leu His Ile Glu  
 2130 2135 2140

Asp Tyr Ser Val Ser Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe  
 2145 2150 2155 2160

Ala Lys Asp Gln Ser Asp Asp Asp His Leu Lys Asp Leu Ser Leu His  
 2165 2170 2175

Lys Asn Gln Thr Val Val Asp Val Ala Val Leu Thr Ser Phe Leu Gln  
 2180 2185 2190

Asp Glu Lys Val Lys Glu Ser Tyr Val  
 2195 2200

<210> 10

<211> 2310

<212> PRT

<213> Mus musculus

<400> 10

Met Gly Phe Leu Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr  
 1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro  
 20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu  
 35 40 45

Tyr Ser Gln His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala  
 50 55 60

Gly Leu Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Met Asn Asn Pro  
 65 70 75 80

Cys Phe Gln Asn Pro Thr Pro Gly Glu Ser Pro Gly Thr Val Ser Asn  
 85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu  
 100 105 110

Phe Met Asp Thr Pro Glu Val Gln His Leu Gly Gln Val Trp Ala Glu  
 115 120 125

Leu Arg Thr Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu  
 130 135 140

Arg Phe Ala Gly Arg Gly Leu Gln Ile Arg Asp Ile Leu Lys Asp Glu  
 145 150 155 160

Glu Ala Leu Thr Leu Phe Leu Met Arg Asn Ile Gly Leu Ser Asp Ser  
 165 170 175  
 Val Ala His Leu Leu Val Asn Ser Gln Val Arg Val Glu Gln Phe Ala  
 180 185 190  
 Tyr Gly Val Pro Asp Leu Glu Leu Thr Asp Ile Ala Cys Ser Glu Ala  
 195 200 205  
 Leu Leu Gln Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Gln Thr  
 210 215 220  
 Val Arg Asp Ala Leu Cys Pro Leu Ser Gln Val Thr Leu Gln Trp Ile  
 225 230 235 240  
 Glu Asp Thr Leu Tyr Ala Asp Val Asp Phe Phe Lys Leu Phe His Val  
 245 250 255  
 Leu Pro Thr Leu Leu Asp Ser Ser Ser Gln Gly Ile Asn Leu Arg Phe  
 260 265 270  
 Trp Gly Gly Ile Leu Ser Asp Leu Ser Pro Arg Met Gln Lys Phe Ile  
 275 280 285  
 His Arg Pro Ser Val Gln Asp Leu Leu Trp Val Ser Arg Pro Leu Leu  
 290 295 300  
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Gln Leu Met Ser Ile Leu Ser  
 305 310 315 320  
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Phe Ser  
 325 330 335  
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp  
 340 345 350  
 Ser Thr Arg Lys Asp Pro Ala Tyr Ser Tyr Asp Lys Arg Thr Thr Ser  
 355 360 365  
 Phe Cys Asn Ser Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys  
 370 375 380  
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Phe  
 385 390 395 400  
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Met Lys Asn Ala Asn Ser  
 405 410 415  
 Thr Phe Glu Glu Leu Asp Arg Val Arg Lys Leu Val Lys Ala Trp Glu  
 420 425 430  
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Glu Lys Ser Thr Gln Met  
 435 440 445  
 Thr Val Ile Arg Asp Thr Leu Gln His Pro Thr Val Lys Asp Phe Ile  
 450 455 460

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Arg | Gln | Leu | Gly | Glu | Glu | Gly | Ile | Thr | Thr | Glu | Ala | Val | Leu | Asn | 465 | 470 | 475 | 480 |
| Phe | Phe | Ser | Asn | Gly | Pro | Gln | Glu | Lys | Gln | Ala | Asp | Asp | Met | Thr | Ser | 485 | 490 |     | 495 |
| Phe | Asp | Trp | Arg | Asp | Ile | Phe | Asn | Ile | Thr | Asp | Arg | Phe | Leu | Arg | Leu | 500 | 505 |     | 510 |
| Ala | Asn | Gln | Tyr | Leu | Glu | Cys | Leu | Val | Leu | Asp | Lys | Phe | Glu | Ser | Tyr | 515 | 520 |     | 525 |
| Asp | Asp | Glu | Val | Gln | Leu | Thr | Gln | Arg | Ala | Leu | Ser | Leu | Leu | Glu | Glu | 530 | 535 | 540 |     |
| Asn | Arg | Phe | Trp | Ala | Gly | Val | Val | Phe | Pro | Gly | Met | Tyr | Pro | Trp | Ala | 545 | 550 | 555 | 560 |
| Ser | Ser | Leu | Pro | Pro | His | Val | Lys | Tyr | Lys | Ile | Arg | Met | Asp | Ile | Asp | 565 | 570 |     | 575 |
| Val | Val | Glu | Lys | Thr | Asn | Lys | Ile | Lys | Asp | Arg | Tyr | Trp | Asp | Ser | Gly | 580 | 585 |     | 590 |
| Pro | Arg | Ala | Asp | Pro | Val | Glu | Asp | Phe | Arg | Tyr | Ile | Trp | Gly | Gly | Phe | 595 | 600 | 605 |     |
| Ala | Tyr | Leu | Gln | Asp | Met | Val | Glu | Gln | Gly | Ile | Val | Lys | Ser | Gln | Met | 610 | 615 | 620 |     |
| Gln | Ala | Glu | Pro | Pro | Ile | Gly | Val | Tyr | Leu | Gln | Gln | Met | Pro | Tyr | Pro | 625 | 630 | 635 | 640 |
| Cys | Phe | Val | Asp | Asp | Ser | Phe | Met | Ile | Ile | Leu | Asn | Arg | Cys | Phe | Pro | 645 | 650 |     | 655 |
| Ile | Phe | Met | Val | Leu | Ala | Trp | Ile | Tyr | Ser | Val | Ser | Met | Thr | Val | Lys | 660 | 665 |     | 670 |
| Gly | Ile | Val | Leu | Glu | Lys | Glu | Leu | Arg | Leu | Lys | Glu | Thr | Leu | Lys | Asn | 675 | 680 | 685 |     |
| Gln | Gly | Val | Ser | Asn | Ala | Val | Ile | Trp | Cys | Thr | Trp | Phe | Leu | Asp | Ser | 690 | 695 | 700 |     |
| Phe | Ser | Ile | Met | Ala | Leu | Ser | Ile | Phe | Leu | Leu | Thr | Leu | Phe | Ile | Met | 705 | 710 | 715 | 720 |
| His | Gly | Arg | Ile | Leu | His | Tyr | Ser | Asp | Pro | Phe | Ile | Leu | Phe | Leu | Phe | 725 | 730 |     | 735 |
| Leu | Leu | Ala | Phe | Ala | Thr | Ala | Thr | Ile | Met | Gln | Ser | Phe | Leu | Leu | Ser | 740 | 745 |     | 750 |
| Thr | Leu | Phe | Ser | Lys | Ala | Ser | Leu | Ala | Ala | Ala | Cys | Ser | Gly | Val | Ile | 755 | 760 | 765 |     |
| Tyr | Phe | Thr | Leu | Tyr | Leu | Pro | His | Val | Leu | Cys | Phe | Ala | Trp | Gln | Asp |     |     |     |     |

| 770  |      |     |      |      | 775  |      |     |      |     | 780  |      |      |      |     |      |
|------|------|-----|------|------|------|------|-----|------|-----|------|------|------|------|-----|------|
| Arg  | Met  | Thr | Ala  | Asp  | Leu  | Lys  | Thr | Thr  | Val | Ser  | Leu  | Leu  | Ser  | Ser | Val  |
| 785  |      |     |      |      | 790  |      |     |      |     | 795  |      |      |      |     | 800  |
| Ala  | Phe  | Gly | Phe  | Gly  | Thr  | Glu  | Tyr | Leu  | Val | Arg  | Phe  | Glu  | Glu  | Gln | Gly  |
|      |      |     |      | 805  |      |      |     |      | 810 |      |      |      |      | 815 |      |
| Leu  | Gly  | Leu | Gln  | Trp  | Ser  | Asn  | Ile | Gly  | Lys | Ser  | Pro  | Leu  | Glu  | Gly | Asp  |
|      |      |     | 820  |      |      |      |     | 825  |     |      |      |      | 830  |     |      |
| Glu  | Phe  | Ser | Phe  | Leu  | Leu  | Ser  | Met | Lys  | Met | Met  | Leu  | Leu  | Asp  | Ala | Ala  |
|      |      | 835 |      |      |      |      | 840 |      |     |      |      | 845  |      |     |      |
| Leu  | Tyr  | Gly | Leu  | Leu  | Ala  | Trp  | Tyr | Leu  | Asp | Gln  | Val  | Phe  | Pro  | Gly | Asp  |
|      | 850  |     |      |      |      | 855  |     |      |     |      | 860  |      |      |     |      |
| Tyr  | Gly  | Thr | Pro  | Leu  | Pro  | Trp  | Tyr | Phe  | Leu | Leu  | Gln  | Glu  | Ser  | Tyr | Trp  |
| 865  |      |     |      |      | 870  |      |     |      |     | 875  |      |      |      |     | 880  |
| Leu  | Gly  | Gly | Glu  | Gly  | Cys  | Ser  | Thr | Arg  | Glu | Glu  | Arg  | Ala  | Leu  | Glu | Lys  |
|      |      |     |      | 885  |      |      |     |      | 890 |      |      |      |      | 895 |      |
| Thr  | Glu  | Pro | Leu  | Thr  | Glu  | Glu  | Met | Glu  | Asp | Pro  | Glu  | His  | Pro  | Glu | Gly  |
|      |      |     | 900  |      |      |      |     | 905  |     |      |      |      | 910  |     |      |
| Met  | Asn  | Asp | Ser  | Phe  | Phe  | Glu  | Arg | Glu  | Leu | Pro  | Gly  | Leu  | Val  | Pro | Gly  |
|      |      | 915 |      |      |      |      | 920 |      |     |      |      | 925  |      |     |      |
| Val  | Cys  | Val | Lys  | Asn  | Leu  | Val  | Lys | Val  | Phe | Glu  | Pro  | Ser  | Gly  | Arg | Pro  |
|      | 930  |     |      |      |      | 935  |     |      |     |      | 940  |      |      |     |      |
| Ala  | Val  | Asp | Arg  | Leu  | Asn  | Ile  | Thr | Phe  | Tyr | Glu  | Asn  | Gln  | Ile  | Thr | Ala  |
| 945  |      |     |      |      | 950  |      |     |      |     | 955  |      |      |      |     | 960  |
| Phe  | Leu  | Gly | His  | Asn  | Gly  | Ala  | Gly | Lys  | Thr | Thr  | Thr  | Leu  | Ser  | Ile | Leu  |
|      |      |     |      | 965  |      |      |     |      | 970 |      |      |      |      | 975 |      |
| Thr  | Gly  | Leu | Leu  | Pro  | Pro  | Thr  | Ser | Gly  | Thr | Val  | Leu  | Ile  | Gly  | Gly | Lys  |
|      |      |     | 980  |      |      |      |     | 985  |     |      |      |      | 990  |     |      |
| Asp  | Ile  | Glu | Thr  | Asn  | Leu  | Asp  | Val | Val  | Arg | Gln  | Ser  | Leu  | Gly  | Met | Cys  |
|      | 995  |     |      |      |      | 1000 |     |      |     |      | 1005 |      |      |     |      |
| Pro  | Gln  | His | Asn  | Ile  | Leu  | Phe  | His | His  | Leu | Thr  | Val  | Ala  | Glu  | His | Ile  |
|      | 1010 |     |      |      |      | 1015 |     |      |     |      | 1020 |      |      |     |      |
| Leu  | Phe  | Tyr | Ala  | Gln  | Leu  | Lys  | Gly | Arg  | Ser | Trp  | Glu  | Glu  | Ala  | Gln | Leu  |
| 1025 |      |     |      |      | 1030 |      |     |      |     | 1035 |      |      |      |     | 1040 |
| Glu  | Met  | Glu | Ala  | Met  | Leu  | Glu  | Asp | Thr  | Gly | Leu  | His  | His  | Lys  | Arg | Asn  |
|      |      |     |      | 1045 |      |      |     | 1050 |     |      |      |      | 1055 |     |      |
| Glu  | Glu  | Ala | Gln  | Asp  | Leu  | Ser  | Gly | Gly  | Met | Gln  | Arg  | Lys  | Leu  | Ser | Val  |
|      |      |     | 1060 |      |      |      |     | 1065 |     |      |      |      | 1070 |     |      |
| Ala  | Ile  | Ala | Phe  | Val  | Gly  | Asp  | Ser | Lys  | Val | Val  | Val  | Leu  | Asp  | Glu | Pro  |
|      | 1075 |     |      |      |      | 1080 |     |      |     |      |      | 1085 |      |     |      |

Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu  
 1090 1095 1100  
 Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met  
 1105 1110 1115 1120  
 Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ser Gln Gly  
 1125 1130 1135  
 Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly  
 1140 1145 1150  
 Thr Gly Phe Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser  
 1155 1160 1165  
 Gln Arg Gly Gly Cys Glu Gly Val Cys Ser Cys Thr Ser Lys Gly Phe  
 1170 1175 1180  
 Ser Thr Arg Cys Pro Thr Arg Val Asp Glu Ile Thr Glu Glu Gln Val  
 1185 1190 1195 1200  
 Leu Asp Gly Asp Val Gln Glu Leu Met Asp Leu Val Tyr His His Val  
 1205 1210 1215  
 Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu  
 1220 1225 1230  
 Leu Pro Asn Lys Asn Phe Lys Gln Arg Ala Tyr Ala Ser Leu Phe Arg  
 1235 1240 1245  
 Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly Ile  
 1250 1255 1260  
 Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp Ala  
 1265 1270 1275 1280  
 Gly Ala Gly Ser Met Phe Val Gly Gly Ala Gln Gln Lys Arg Glu Gln  
 1285 1290 1295  
 Ala Gly Leu Arg His Pro Cys Ser Ala Pro Thr Glu Lys Leu Arg Gln  
 1300 1305 1310  
 Tyr Ala Gln Ala Pro His Thr Cys Ser Pro Gly Gln Val Asp Pro Pro  
 1315 1320 1325  
 Lys Gly Gln Pro Ser Pro Glu Pro Glu Asp Pro Gly Val Pro Phe Asn  
 1330 1335 1340  
 Thr Gly Ala Arg Leu Ile Leu Gln His Val Gln Ala Leu Leu Val Lys  
 1345 1350 1355 1360  
 Arg Phe His His Thr Ile Arg Ser Arg Lys Asp Phe Val Ala Gln Ile  
 1365 1370 1375  
 Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu Met Leu Ser Ile Ile  
 1380 1385 1390

Val Pro Pro Phe Gly Glu Phe Pro Ala Leu Thr Leu His Pro Trp Met  
 1395 1400 1405  
 Tyr Gly His Gln Tyr Thr Phe Phe Ser Met Asp Glu Pro Asn Asn Glu  
 1410 1415 1420  
 His Leu Glu Val Leu Ala Asp Val Leu Leu Asn Arg Pro Gly Phe Gly  
 1425 1430 1435 1440  
 Asn Arg Cys Leu Lys Glu Glu Trp Leu Pro Glu Tyr Pro Cys Ile Asn  
 1445 1450 1455  
 Ala Thr Ser Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr His Leu  
 1460 1465 1470  
 Phe Gln Lys Gln Lys Trp Thr Ala Ala His Pro Ser Pro Ser Cys Lys  
 1475 1480 1485  
 Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro Glu Gly  
 1490 1495 1500  
 Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser Thr Glu Val  
 1505 1510 1515 1520  
 Leu Gln Asp Leu Thr Asn Arg Asn Ile Ser Asp Tyr Leu Val Lys Thr  
 1525 1530 1535  
 Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser Lys Phe Trp Val Asn  
 1540 1545 1550  
 Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly Lys Leu Pro Ala Ile  
 1555 1560 1565  
 Pro Ile Ser Gly Glu Ala Leu Val Gly Phe Leu Ser Gly Leu Gly Gln  
 1570 1575 1580  
 Met Met Asn Val Ser Gly Gly Pro Val Thr Arg Glu Ala Ser Lys Glu  
 1585 1590 1595 1600  
 Met Leu Asp Phe Leu Lys His Leu Glu Thr Thr Asp Asn Ile Lys Val  
 1605 1610 1615  
 Trp Phe Asn Asn Lys Gly Trp His Ala Leu Val Ser Phe Leu Asn Val  
 1620 1625 1630  
 Ala His Asn Ala Ile Leu Arg Ala Ser Leu Pro Arg Asp Arg Asp Pro  
 1635 1640 1645  
 Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro Leu Asn Leu Thr Lys  
 1650 1655 1660  
 Glu Gln Leu Ser Asp Ile Thr Val Leu Thr Thr Ser Val Asp Ala Val  
 1665 1670 1675 1680  
 Val Ala Ile Cys Val Ile Phe Ala Met Ser Phe Val Pro Ala Ser Phe  
 1685 1690 1695  
 Val Leu Tyr Leu Ile Gln Glu Arg Val Thr Lys Ala Lys His Leu Gln

| 1700                                                                                   | 1705 | 1710 |
|----------------------------------------------------------------------------------------|------|------|
| Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp Leu Thr Asn Phe Leu<br>1715 1720 1725      |      |      |
| Trp Asp Ile Met Asn Tyr Ala Val Ser Ala Gly Leu Val Val Gly Ile<br>1730 1735 1740      |      |      |
| Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Asp Asn Leu Pro<br>1745 1750 1755 1760 |      |      |
| Ala Leu Val Ser Leu Leu Met Leu Tyr Gly Trp Ala Val Ile Pro Met<br>1765 1770 1775      |      |      |
| Met Tyr Pro Ala Ser Phe Leu Phe Glu Val Pro Ser Thr Ala Tyr Val<br>1780 1785 1790      |      |      |
| Ala Leu Ser Cys Ala Asn Leu Phe Ile Gly Ile Asn Ser Ser Ala Ile<br>1795 1800 1805      |      |      |
| Thr Phe Val Leu Glu Leu Phe Glu Asn Asn Arg Thr Leu Leu Arg Phe<br>1810 1815 1820      |      |      |
| Asn Ala Met Leu Arg Lys Leu Leu Ile Val Phe Pro His Phe Cys Leu<br>1825 1830 1835 1840 |      |      |
| Gly Arg Gly Leu Ile Asp Leu Ala Leu Ser Gln Ala Val Thr Asp Val<br>1845 1850 1855      |      |      |
| Tyr Ala Gln Phe Gly Glu Glu Tyr Ser Ala Asn Pro Phe Gln Trp Asp<br>1860 1865 1870      |      |      |
| Leu Ile Gly Lys Asn Leu Val Ala Met Ala Ile Glu Gly Val Val Tyr<br>1875 1880 1885      |      |      |
| Phe Leu Leu Thr Leu Leu Ile Gln His His Phe Phe Leu Thr Arg Trp<br>1890 1895 1900      |      |      |
| Ile Ala Glu Pro Ala Arg Glu Pro Val Phe Asp Glu Asp Asp Asp Val<br>1905 1910 1915 1920 |      |      |
| Ala Glu Glu Arg Gln Arg Val Met Ser Gly Gly Asn Lys Thr Asp Ile<br>1925 1930 1935      |      |      |
| Leu Lys Leu Asn Glu Leu Thr Lys Val Tyr Ser Gly Ser Ser Ser Pro<br>1940 1945 1950      |      |      |
| Ala Val Asp Arg Leu Cys Val Gly Val Arg Pro Gly Glu Cys Phe Gly<br>1955 1960 1965      |      |      |
| Leu Leu Gly Val Asn Gly Ala Gly Lys Thr Thr Thr Phe Lys Met Leu<br>1970 1975 1980      |      |      |
| Thr Gly Asp Thr Thr Val Thr Ser Gly Asp Ala Thr Val Ala Gly Lys<br>1985 1990 1995 2000 |      |      |
| Ser Ile Leu Thr Ser Ile Ser Asp Val His Gln Asn Met Gly Tyr Cys<br>2005 2010 2015      |      |      |

Pro Gln Phe Asp Ala Ile Asp Asp Leu Leu Thr Gly Arg Glu His Leu  
 2020 2025 2030  
 Tyr Leu Tyr Ala Arg Leu Arg Gly Val Pro Ser Lys Glu Ile Glu Lys  
 2035 2040 2045  
 Val Ala Asn Trp Gly Ile Gln Ser Leu Gly Leu Ser Leu Tyr Ala Asp  
 2050 2055 2060  
 Arg Leu Ala Gly Thr Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser Thr  
 2065 2070 2075 2080  
 Ala Ile Ala Leu Thr Gly Cys Pro Pro Leu Leu Leu Leu Asp Glu Pro  
 2085 2090 2095  
 Thr Thr Gly Met Asp Pro Gln Ala Arg Arg Met Leu Trp Asn Thr Ile  
 2100 2105 2110  
 Val Ser Ile Ile Arg Glu Gly Arg Ala Val Val Leu Thr Ser His Ser  
 2115 2120 2125  
 Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met Val Lys  
 2130 2135 2140  
 Gly Thr Phe Gln Cys Leu Gly Thr Ile Gln His Leu Lys Tyr Lys Phe  
 2145 2150 2155 2160  
 Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys Asp Asp  
 2165 2170 2175  
 Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly Asn Phe  
 2180 2185 2190  
 Pro Gly Ser Val Gln Arg Glu Arg His His Ser Met Leu Gln Phe Gln  
 2195 2200 2205  
 Val Pro Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Ile Ser His  
 2210 2215 2220  
 Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr Thr Leu  
 2225 2230 2235 2240  
 Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Thr Tyr Asp  
 2245 2250 2255  
 Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser Trp Gln Ala Lys Leu  
 2260 2265 2270  
 Glu Glu Lys Ser Gly Arg Leu Gln Thr Gln Glu Pro Leu Pro Ala Gly  
 2275 2280 2285  
 Ser Glu Gln Leu Ala Asn Gly Ser Asn Pro Thr Ala Ala Glu Asp Lys  
 2290 2295 2300  
 His Thr Arg Ser Pro Gln  
 2305 2310

<210> 11  
 <211> 2261  
 <212> PRT  
 <213> Homo sapiens  
  
 <400> 11  
 Met Ala Cys Trp Pro Gln Leu Arg Leu Leu Leu Trp Lys Asn Leu Thr  
   1                  5                  10                  15  
  
 Phe Arg Arg Arg Gln Thr Cys Gln Leu Leu Leu Glu Val Ala Trp Pro  
           20                  25                  30  
  
 Leu Phe Ile Phe Leu Ile Leu Ile Ser Val Arg Leu Ser Tyr Pro Pro  
       35                  40                  45  
  
 Tyr Glu Gln His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala  
       50                  55                  60  
  
 Gly Thr Leu Pro Trp Val Gln Gly Ile Ile Cys Asn Ala Asn Asn Pro  
       65                  70                  75                  80  
  
 Cys Phe Arg Tyr Pro Thr Pro Gly Glu Ala Pro Gly Val Val Gly Asn  
           85                  90                  95  
  
 Phe Asn Lys Ser Ile Val Ala Arg Leu Phe Ser Asp Ala Arg Arg Leu  
          100                 105                 110  
  
 Leu Leu Tyr Ser Gln Lys Asp Thr Ser Met Lys Asp Met Arg Lys Val  
      115                 120                 125  
  
 Leu Arg Thr Leu Gln Gln Ile Lys Lys Ser Ser Ser Asn Leu Lys Leu  
      130                 135                 140  
  
 Gln Asp Phe Leu Val Asp Asn Glu Thr Phe Ser Gly Phe Leu Tyr His  
   145                 150                 155                 160  
  
 Asn Leu Ser Leu Pro Lys Ser Thr Val Asp Lys Met Leu Arg Ala Asp  
          165                 170                 175  
  
 Val Ile Leu His Lys Val Phe Leu Gln Gly Tyr Gln Leu His Leu Thr  
      180                 185                 190  
  
 Ser Leu Cys Asn Gly Ser Lys Ser Glu Glu Met Ile Gln Leu Gly Asp  
      195                 200                 205  
  
 Gln Glu Val Ser Glu Leu Cys Gly Leu Pro Arg Glu Lys Leu Ala Ala  
      210                 215                 220  
  
 Ala Glu Arg Val Leu Arg Ser Asn Met Asp Ile Leu Lys Pro Ile Leu  
   225                 230                 235                 240  
  
 Arg Thr Leu Asn Ser Thr Ser Pro Phe Pro Ser Lys Glu Leu Ala Glu  
          245                 250                 255  
  
 Ala Thr Lys Thr Leu Leu His Ser Leu Gly Thr Leu Ala Gln Glu Leu  
      260                 265                 270

Phe Ser Met Arg Ser Trp Ser Asp Met Arg Gln Glu Val Met Phe Leu  
 275 280 285  
 Thr Asn Val Asn Ser Ser Ser Ser Thr Gln Ile Tyr Gln Ala Val  
 290 295 300  
 Ser Arg Ile Val Cys Gly His Pro Glu Gly Gly Gly Leu Lys Ile Lys  
 305 310 315 320  
 Ser Leu Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Leu Phe Gly Gly  
 325 330 335  
 Asn Gly Thr Glu Glu Asp Ala Glu Thr Phe Tyr Asp Asn Ser Thr Thr  
 340 345 350  
 Pro Tyr Cys Asn Asp Leu Met Lys Asn Leu Glu Ser Ser Pro Leu Ser  
 355 360 365  
 Arg Ile Ile Trp Lys Ala Leu Lys Pro Leu Leu Val Gly Lys Ile Leu  
 370 375 380  
 Tyr Thr Pro Asp Thr Pro Ala Thr Arg Gln Val Met Ala Glu Val Asn  
 385 390 395 400  
 Lys Thr Phe Gln Glu Leu Ala Val Phe His Asp Leu Glu Gly Met Trp  
 405 410 415  
 Glu Glu Leu Ser Pro Lys Ile Trp Thr Phe Met Glu Asn Ser Gln Glu  
 420 425 430  
 Met Asp Leu Val Arg Met Leu Leu Asp Ser Arg Asp Asn Asp His Phe  
 435 440 445  
 Trp Glu Gln Gln Leu Asp Gly Leu Asp Trp Thr Ala Gln Asp Ile Val  
 450 455 460  
 Ala Phe Leu Ala Lys His Pro Glu Asp Val Gln Ser Ser Asn Gly Ser  
 465 470 475 480  
 Val Tyr Thr Trp Arg Glu Ala Phe Asn Glu Thr Asn Gln Ala Ile Arg  
 485 490 495  
 Thr Ile Ser Arg Phe Met Glu Cys Val Asn Leu Asn Lys Leu Glu Pro  
 500 505 510  
 Ile Ala Thr Glu Val Trp Leu Ile Asn Lys Ser Met Glu Leu Leu Asp  
 515 520 525  
 Glu Arg Lys Phe Trp Ala Gly Ile Val Phe Thr Gly Ile Thr Pro Gly  
 530 535 540  
 Ser Ile Glu Leu Pro His His Val Lys Tyr Lys Ile Arg Met Asp Ile  
 545 550 555 560  
 Asp Asn Val Glu Arg Thr Asn Lys Ile Lys Asp Gly Tyr Trp Asp Pro  
 565 570 575  
 Gly Pro Arg Ala Asp Pro Phe Glu Asp Met Arg Tyr Val Trp Gly Gly

| 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Phe | Ala | Tyr | Leu | Gln | Asp | Val | Val | Glu | Gln | Ala | Ile | Ile | Arg | Val | Leu |
|     |     | 595 |     |     |     |     |     | 600 |     |     |     |     | 605 |     |     |
| Thr | Gly | Thr | Glu | Lys | Lys | Thr | Gly | Val | Tyr | Met | Gln | Gln | Met | Pro | Tyr |
|     | 610 |     |     |     |     | 615 |     |     |     |     | 620 |     |     |     |     |
| Pro | Cys | Tyr | Val | Asp | Asp | Ile | Phe | Leu | Arg | Val | Met | Ser | Arg | Ser | Met |
| 625 |     |     |     |     | 630 |     |     |     |     | 635 |     |     |     |     | 640 |
| Pro | Leu | Phe | Met | Thr | Leu | Ala | Trp | Ile | Tyr | Ser | Val | Ala | Val | Ile | Ile |
|     |     |     |     | 645 |     |     |     |     | 650 |     |     |     |     | 655 |     |
| Lys | Gly | Ile | Val | Tyr | Glu | Lys | Glu | Ala | Arg | Leu | Lys | Glu | Thr | Met | Arg |
|     |     |     | 660 |     |     |     |     | 665 |     |     |     |     | 670 |     |     |
| Ile | Met | Gly | Leu | Asp | Asn | Ser | Ile | Leu | Trp | Phe | Ser | Trp | Phe | Ile | Ser |
|     |     | 675 |     |     |     |     | 680 |     |     |     |     | 685 |     |     |     |
| Ser | Leu | Ile | Pro | Leu | Leu | Val | Ser | Ala | Gly | Leu | Leu | Val | Val | Ile | Leu |
|     | 690 |     |     |     |     | 695 |     |     |     |     | 700 |     |     |     |     |
| Lys | Leu | Gly | Asn | Leu | Leu | Pro | Tyr | Ser | Asp | Pro | Ser | Val | Val | Phe | Val |
| 705 |     |     |     |     | 710 |     |     |     |     | 715 |     |     |     |     | 720 |
| Phe | Leu | Ser | Val | Phe | Ala | Val | Val | Thr | Ile | Leu | Gln | Cys | Phe | Leu | Ile |
|     |     |     |     | 725 |     |     |     |     | 730 |     |     |     |     | 735 |     |
| Ser | Thr | Leu | Phe | Ser | Arg | Ala | Asn | Leu | Ala | Ala | Ala | Cys | Gly | Gly | Ile |
|     |     |     | 740 |     |     |     |     | 745 |     |     |     |     | 750 |     |     |
| Ile | Tyr | Phe | Thr | Leu | Tyr | Leu | Pro | Tyr | Val | Leu | Cys | Val | Ala | Trp | Gln |
|     |     | 755 |     |     |     |     | 760 |     |     |     |     | 765 |     |     |     |
| Asp | Tyr | Val | Gly | Phe | Thr | Leu | Lys | Ile | Phe | Ala | Ser | Leu | Leu | Ser | Pro |
|     | 770 |     |     |     |     | 775 |     |     |     |     | 780 |     |     |     |     |
| Val | Ala | Phe | Gly | Phe | Gly | Cys | Glu | Tyr | Phe | Ala | Leu | Phe | Glu | Glu | Gln |
| 785 |     |     |     |     | 790 |     |     |     |     | 795 |     |     |     |     | 800 |
| Gly | Ile | Gly | Val | Gln | Trp | Asp | Asn | Leu | Phe | Glu | Ser | Pro | Val | Glu | Glu |
|     |     |     |     | 805 |     |     |     |     | 810 |     |     |     |     | 815 |     |
| Asp | Gly | Phe | Asn | Leu | Thr | Thr | Ser | Val | Ser | Met | Met | Leu | Phe | Asp | Thr |
|     |     |     | 820 |     |     |     |     | 825 |     |     |     |     | 830 |     |     |
| Phe | Leu | Tyr | Gly | Val | Met | Thr | Trp | Tyr | Ile | Glu | Ala | Val | Phe | Pro | Gly |
|     |     | 835 |     |     |     |     | 840 |     |     |     |     | 845 |     |     |     |
| Gln | Tyr | Gly | Ile | Pro | Arg | Pro | Trp | Tyr | Phe | Pro | Cys | Thr | Lys | Ser | Tyr |
|     | 850 |     |     |     |     | 855 |     |     |     |     | 860 |     |     |     |     |
| Trp | Phe | Gly | Glu | Glu | Ser | Asp | Glu | Lys | Ser | His | Pro | Gly | Ser | Asn | Gln |
| 865 |     |     |     |     | 870 |     |     |     |     | 875 |     |     |     |     | 880 |
| Lys | Arg | Ile | Ser | Glu | Ile | Cys | Met | Glu | Glu | Glu | Pro | Thr | His | Leu | Lys |
|     |     |     |     | 885 |     |     |     |     | 890 |     |     |     |     | 895 |     |

Leu Gly Val Ser Ile Gln Asn Leu Val Lys Val Tyr Arg Asp Gly Met  
 900 905 910  
 Lys Val Ala Val Asp Gly Leu Ala Leu Asn Phe Tyr Glu Gly Gln Ile  
 915 920 925  
 Thr Ser Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Met Ser  
 930 935 940  
 Ile Leu Thr Gly Leu Phe Pro Pro Thr Ser Gly Thr Ala Tyr Ile Leu  
 945 950 955 960  
 Gly Lys Asp Ile Arg Ser Glu Met Ser Thr Ile Arg Gln Asn Leu Gly  
 965 970 975  
 Val Cys Pro Gln His Asn Val Leu Phe Asp Met Leu Thr Val Glu Glu  
 980 985 990  
 His Ile Trp Phe Tyr Ala Arg Leu Lys Gly Leu Ser Glu Lys His Val  
 995 1000 1005  
 Lys Ala Glu Met Glu Gln Met Ala Leu Asp Val Gly Leu Pro Ser Ser  
 1010 1015 1020  
 Lys Leu Lys Ser Lys Thr Ser Gln Leu Ser Gly Gly Met Gln Arg Lys  
 1025 1030 1035 1040  
 Leu Ser Val Ala Leu Ala Phe Val Gly Gly Ser Lys Val Val Ile Leu  
 1045 1050 1055  
 Asp Glu Pro Thr Ala Gly Val Asp Pro Tyr Ser Arg Arg Gly Ile Trp  
 1060 1065 1070  
 Glu Leu Leu Leu Lys Tyr Arg Gln Gly Arg Thr Ile Ile Leu Ser Thr  
 1075 1080 1085  
 His His Met Asp Glu Ala Asp Val Leu Gly Asp Arg Ile Ala Ile Ile  
 1090 1095 1100  
 Ser His Gly Lys Leu Cys Cys Val Gly Ser Ser Leu Phe Leu Lys Asn  
 1105 1110 1115 1120  
 Gln Leu Gly Thr Gly Tyr Tyr Leu Thr Leu Val Lys Lys Asp Val Glu  
 1125 1130 1135  
 Ser Ser Leu Ser Ser Cys Arg Asn Ser Ser Ser Thr Val Ser Tyr Leu  
 1140 1145 1150  
 Lys Lys Glu Asp Ser Val Ser Gln Ser Ser Ser Asp Ala Gly Leu Gly  
 1155 1160 1165  
 Ser Asp His Glu Ser Asp Thr Leu Thr Ile Asp Val Ser Ala Ile Ser  
 1170 1175 1180  
 Asn Leu Ile Arg Lys His Val Ser Glu Ala Arg Leu Val Glu Asp Ile  
 1185 1190 1195 1200

Gly His Glu Leu Thr Tyr Val Leu Pro Tyr Glu Ala Ala Lys Glu Gly  
 1205 1210 1215  
 Ala Phe Val Glu Leu Phe His Glu Ile Asp Asp Arg Leu Ser Asp Leu  
 1220 1225 1230  
 Gly Ile Ser Ser Tyr Gly Ile Ser Glu Thr Thr Leu Glu Glu Ile Phe  
 1235 1240 1245  
 Leu Lys Val Ala Glu Glu Ser Gly Val Asp Ala Glu Thr Ser Asp Gly  
 1250 1255 1260  
 Thr Leu Pro Ala Arg Arg Asn Arg Arg Ala Phe Gly Asp Lys Gln Ser  
 1265 1270 1275 1280  
 Cys Leu Arg Pro Phe Thr Glu Asp Asp Ala Ala Asp Pro Asn Asp Ser  
 1285 1290 1295  
 Asp Ile Asp Pro Glu Ser Arg Glu Thr Asp Leu Leu Ser Gly Met Asp  
 1300 1305 1310  
 Gly Lys Gly Ser Tyr Gln Val Lys Gly Trp Lys Leu Thr Gln Gln Gln  
 1315 1320 1325  
 Phe Val Ala Leu Leu Trp Lys Arg Leu Leu Ile Ala Arg Arg Ser Arg  
 1330 1335 1340  
 Lys Gly Phe Phe Ala Gln Ile Val Leu Pro Ala Val Phe Val Cys Ile  
 1345 1350 1355 1360  
 Ala Leu Val Phe Ser Leu Ile Val Pro Pro Phe Gly Lys Tyr Pro Ser  
 1365 1370 1375  
 Leu Glu Leu Gln Pro Trp Met Tyr Asn Glu Gln Tyr Thr Phe Val Ser  
 1380 1385 1390  
 Asn Asp Ala Pro Glu Asp Thr Gly Thr Leu Glu Leu Leu Asn Ala Leu  
 1395 1400 1405  
 Thr Lys Asp Pro Gly Phe Gly Thr Arg Cys Met Glu Gly Asn Pro Ile  
 1410 1415 1420  
 Pro Asp Thr Pro Cys Gln Ala Gly Glu Glu Glu Trp Thr Thr Ala Pro  
 1425 1430 1435 1440  
 Val Pro Gln Thr Ile Met Asp Leu Phe Gln Asn Gly Asn Trp Thr Met  
 1445 1450 1455  
 Gln Asn Pro Ser Pro Ala Cys Gln Cys Ser Ser Asp Lys Ile Lys Lys  
 1460 1465 1470  
 Met Leu Pro Val Cys Pro Pro Gly Ala Gly Gly Leu Pro Pro Gln  
 1475 1480 1485  
 Arg Lys Gln Asn Thr Ala Asp Ile Leu Gln Asp Leu Thr Gly Arg Asn  
 1490 1495 1500  
 Ile Ser Asp Tyr Leu Val Lys Thr Tyr Val Gln Ile Ile Ala Lys Ser

|                                                                 |      |      |      |
|-----------------------------------------------------------------|------|------|------|
| 1505                                                            | 1510 | 1515 | 1520 |
| Leu Lys Asn Lys Ile Trp Val Asn Glu Phe Arg Tyr Gly Gly Phe Ser | 1525 | 1530 | 1535 |
| Leu Gly Val Ser Asn Thr Gln Ala Leu Pro Pro Ser Gln Glu Val Asn | 1540 | 1545 | 1550 |
| Asp Ala Thr Lys Gln Met Lys Lys His Leu Lys Leu Ala Lys Asp Ser | 1555 | 1560 | 1565 |
| Ser Ala Asp Arg Phe Leu Asn Ser Leu Gly Arg Phe Met Thr Gly Leu | 1570 | 1575 | 1580 |
| Asp Thr Arg Asn Asn Val Lys Val Trp Phe Asn Asn Lys Gly Trp His | 1585 | 1590 | 1595 |
| Ala Ile Ser Ser Phe Leu Asn Val Ile Asn Asn Ala Ile Leu Arg Ala | 1605 | 1610 | 1615 |
| Asn Leu Gln Lys Gly Glu Asn Pro Ser His Tyr Gly Ile Thr Ala Phe | 1620 | 1625 | 1630 |
| Asn His Pro Leu Asn Leu Thr Lys Gln Gln Leu Ser Glu Val Ala Pro | 1635 | 1640 | 1645 |
| Met Thr Thr Ser Val Asp Val Leu Val Ser Ile Cys Val Ile Phe Ala | 1650 | 1655 | 1660 |
| Met Ser Phe Val Pro Ala Ser Phe Val Val Phe Leu Ile Gln Glu Arg | 1665 | 1670 | 1675 |
| Val Ser Lys Ala Lys His Leu Gln Phe Ile Ser Gly Val Lys Pro Val | 1685 | 1690 | 1695 |
| Ile Tyr Trp Leu Ser Asn Phe Val Trp Asp Met Cys Asn Tyr Val Val | 1700 | 1705 | 1710 |
| Pro Ala Thr Leu Val Ile Ile Ile Phe Ile Cys Phe Gln Gln Lys Ser | 1715 | 1720 | 1725 |
| Tyr Val Ser Ser Thr Asn Leu Pro Val Leu Ala Leu Leu Leu Leu     | 1730 | 1735 | 1740 |
| Tyr Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala Ser Phe Val Phe | 1745 | 1750 | 1755 |
| Lys Ile Pro Ser Thr Ala Tyr Val Val Leu Thr Ser Val Asn Leu Phe | 1765 | 1770 | 1775 |
| Ile Gly Ile Asn Gly Ser Val Ala Thr Phe Val Leu Glu Leu Phe Thr | 1780 | 1785 | 1790 |
| Asp Asn Lys Leu Asn Asn Ile Asn Asp Ile Leu Lys Ser Val Phe Leu | 1795 | 1800 | 1805 |
| Ile Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Met Val Lys | 1810 | 1815 | 1820 |

Asn Gln Ala Met Ala Asp Ala Leu Glu Arg Phe Gly Glu Asn Arg Phe  
 1825 1830 1835 1840  
 Val Ser Pro Leu Ser Trp Asp Leu Val Gly Arg Asn Leu Phe Ala Met  
 1845 1850 1855  
 Ala Val Glu Gly Val Val Phe Phe Leu Ile Thr Val Leu Ile Gln Tyr  
 1860 1865 1870  
 Arg Phe Phe Ile Arg Pro Arg Pro Val Asn Ala Lys Leu Ser Pro Leu  
 1875 1880 1885  
 Asn Asp Glu Asp Glu Asp Val Arg Arg Glu Arg Gln Arg Ile Leu Asp  
 1890 1895 1900  
 Gly Gly Gly Gln Asn Asp Ile Leu Glu Ile Lys Glu Leu Thr Lys Ile  
 1905 1910 1915 1920  
 Tyr Arg Arg Lys Arg Lys Pro Ala Val Asp Arg Ile Cys Val Gly Ile  
 1925 1930 1935  
 Pro Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys  
 1940 1945 1950  
 Ser Ser Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Arg Gly  
 1955 1960 1965  
 Asp Ala Phe Leu Asn Arg Asn Ser Ile Leu Ser Asn Ile His Glu Val  
 1970 1975 1980  
 His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Thr Glu Leu  
 1985 1990 1995 2000  
 Leu Thr Gly Arg Glu His Val Glu Phe Phe Ala Leu Leu Arg Gly Val  
 2005 2010 2015  
 Pro Glu Lys Glu Val Gly Lys Val Gly Glu Trp Ala Ile Arg Lys Leu  
 2020 2025 2030  
 Gly Leu Val Lys Tyr Gly Glu Lys Tyr Ala Gly Asn Tyr Ser Gly Gly  
 2035 2040 2045  
 Asn Lys Arg Lys Leu Ser Thr Ala Met Ala Leu Ile Gly Gly Pro Pro  
 2050 2055 2060  
 Val Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Lys Ala Arg  
 2065 2070 2075 2080  
 Arg Phe Leu Trp Asn Cys Ala Leu Ser Val Val Lys Glu Gly Arg Ser  
 2085 2090 2095  
 Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr  
 2100 2105 2110  
 Arg Met Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu Gly Ser Val  
 2115 2120 2125

Gln His Leu Lys Asn Arg Phe Gly Asp Gly Tyr Thr Ile Val Val Arg  
 2130 2135 2140  
 Ile Ala Gly Ser Asn Pro Asp Leu Lys Pro Val Gln Asp Phe Phe Gly  
 2145 2150 2155 2160  
 Leu Ala Phe Pro Gly Ser Val Pro Lys Glu Lys His Arg Asn Met Leu  
 2165 2170 2175  
 Gln Tyr Gln Leu Pro Ser Ser Leu Ser Ser Leu Ala Arg Ile Phe Ser  
 2180 2185 2190  
 Ile Leu Ser Gln Ser Lys Lys Arg Leu His Ile Glu Asp Tyr Ser Val  
 2195 2200 2205  
 Ser Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Asp Gln  
 2210 2215 2220  
 Ser Asp Asp Asp His Leu Lys Asp Leu Ser Leu His Lys Asn Gln Thr  
 2225 2230 2235 2240  
 Val Val Asp Val Ala Val Leu Thr Ser Phe Leu Gln Asp Glu Lys Val  
 2245 2250 2255  
 Lys Glu Ser Tyr Val  
 2260

<210> 12  
 <211> 2273  
 <212> PRT  
 <213> Homo sapiens

<400> 12  
 Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr  
 1 5 10 15  
 Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro  
 20 25 30  
 Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu  
 35 40 45  
 Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala  
 50 55 60  
 Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro  
 65 70 75 80  
 Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn  
 85 90 95  
 Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu  
 100 105 110  
 Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu  
 115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu  
 130 135 140  
 Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu  
 145 150 155 160  
 Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser  
 165 170 175  
 Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala  
 180 185 190  
 His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala  
 195 200 205  
 Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr  
 210 215 220  
 Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile  
 225 230 235 240  
 Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val  
 245 250 255  
 Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser  
 260 265 270  
 Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile  
 275 280 285  
 His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met  
 290 295 300  
 Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser  
 305 310 315 320  
 Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser  
 325 330 335  
 Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp  
 340 345 350  
 Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser  
 355 360 365  
 Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys  
 370 375 380  
 Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr  
 385 390 395 400  
 Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser  
 405 410 415  
 Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu  
 420 425 430  
 Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met

| 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Met | Ile | Arg | Asp | Thr | Leu | Gly | Asn | Pro | Thr | Val | Lys | Asp | Phe | Leu |
| 450 |     |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |
| Asn | Arg | Gln | Leu | Gly | Glu | Glu | Gly | Ile | Thr | Ala | Glu | Ala | Ile | Leu | Asn |
| 465 |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |
| Phe | Leu | Tyr | Lys | Gly | Pro | Arg | Glu | Ser | Gln | Ala | Asp | Asp | Met | Ala | Asn |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     | 495 |     |
| Phe | Asp | Trp | Arg | Asp | Ile | Phe | Asn | Ile | Thr | Asp | Arg | Thr | Leu | Arg | Leu |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |     |     |
| Val | Asn | Gln | Tyr | Leu | Glu | Cys | Leu | Val | Leu | Asp | Lys | Phe | Glu | Ser | Tyr |
|     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |     |
| Asn | Asp | Glu | Thr | Gln | Leu | Thr | Gln | Arg | Ala | Leu | Ser | Leu | Leu | Glu | Glu |
|     | 530 |     |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |
| Asn | Met | Phe | Trp | Ala | Gly | Val | Val | Phe | Pro | Asp | Met | Tyr | Pro | Trp | Thr |
| 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |
| Ser | Ser | Leu | Pro | Pro | His | Val | Lys | Tyr | Lys | Ile | Arg | Met | Asp | Ile | Asp |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |
| Val | Val | Glu | Lys | Thr | Asn | Lys | Ile | Lys | Asp | Arg | Tyr | Trp | Asp | Ser | Gly |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |
| Pro | Arg | Ala | Asp | Pro | Val | Glu | Asp | Phe | Arg | Tyr | Ile | Trp | Gly | Gly | Phe |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |
| Ala | Tyr | Leu | Gln | Asp | Met | Val | Glu | Gln | Gly | Ile | Thr | Arg | Ser | Gln | Val |
|     | 610 |     |     |     |     | 615 |     |     |     |     | 620 |     |     |     |     |
| Gln | Ala | Glu | Ala | Pro | Val | Gly | Ile | Tyr | Leu | Gln | Gln | Met | Pro | Tyr | Pro |
| 625 |     |     |     |     | 630 |     |     |     |     | 635 |     |     |     |     | 640 |
| Cys | Phe | Val | Asp | Asp | Ser | Phe | Met | Ile | Ile | Leu | Asn | Arg | Cys | Phe | Pro |
|     |     |     |     | 645 |     |     |     |     | 650 |     |     |     |     | 655 |     |
| Ile | Phe | Met | Val | Leu | Ala | Trp | Ile | Tyr | Ser | Val | Ser | Met | Thr | Val | Lys |
|     |     |     | 660 |     |     |     |     | 665 |     |     |     |     | 670 |     |     |
| Ser | Ile | Val | Leu | Glu | Lys | Glu | Leu | Arg | Leu | Lys | Glu | Thr | Leu | Lys | Asn |
|     |     | 675 |     |     |     |     | 680 |     |     |     |     | 685 |     |     |     |
| Gln | Gly | Val | Ser | Asn | Ala | Val | Ile | Trp | Cys | Thr | Trp | Phe | Leu | Asp | Ser |
|     | 690 |     |     |     |     | 695 |     |     |     |     | 700 |     |     |     |     |
| Phe | Ser | Ile | Met | Ser | Met | Ser | Ile | Phe | Leu | Leu | Thr | Ile | Phe | Ile | Met |
| 705 |     |     |     |     | 710 |     |     |     |     | 715 |     |     |     |     | 720 |
| His | Gly | Arg | Ile | Leu | His | Tyr | Ser | Asp | Pro | Phe | Ile | Leu | Phe | Leu | Phe |
|     |     |     |     | 725 |     |     |     |     | 730 |     |     |     |     | 735 |     |
| Leu | Leu | Ala | Phe | Ser | Thr | Ala | Thr | Ile | Met | Leu | Cys | Phe | Leu | Leu | Ser |
|     |     |     | 740 |     |     |     |     | 745 |     |     |     |     | 750 |     |     |

--

Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile  
 755 760 765  
 Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp  
 770 775 780  
 Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val  
 785 790 795 800  
 Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly  
 805 810 815  
 Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp  
 820 825 830  
 Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala  
 835 840 845  
 Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp  
 850 855 860  
 Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp  
 865 870 875 880  
 Leu Ser Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys  
 885 890 895  
 Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly  
 900 905 910  
 Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly  
 915 920 925  
 Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro  
 930 935 940  
 Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala  
 945 950 955 960  
 Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu  
 965 970 975  
 Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg  
 980 985 990  
 Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys  
 995 1000 1005  
 Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His Met  
 1010 1015 1020  
 Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala Gln Leu  
 1025 1030 1035 1040  
 Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His Lys Arg Asn  
 1045 1050 1055

Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg Lys Leu Ser Val  
 1060 1065 1070  
 Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val Ile Leu Asp Glu Pro  
 1075 1080 1085  
 Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg Ser Ile Trp Asp Leu Leu  
 1090 1095 1100  
 Leu Lys Tyr Arg Ser Gly Arg Thr Ile Ile Met Ser Thr His His Met  
 1105 1110 1115 1120  
 Asp Glu Ala Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ala Gln Gly  
 1125 1130 1135  
 Arg Leu Tyr Cys Ser Gly Thr Pro Leu Phe Leu Lys Asn Cys Phe Gly  
 1140 1145 1150  
 Thr Gly Leu Tyr Leu Thr Leu Val Arg Lys Met Lys Asn Ile Gln Ser  
 1155 1160 1165  
 Gln Arg Lys Gly Ser Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe  
 1170 1175 1180  
 Ser Thr Thr Cys Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln Val  
 1185 1190 1195 1200  
 Leu Asp Gly Asp Val Asn Glu Leu Met Asp Val Val Leu His His Val  
 1205 1210 1215  
 Pro Glu Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu  
 1220 1225 1230  
 Leu Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe Arg  
 1235 1240 1245  
 Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly Ile  
 1250 1255 1260  
 Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu Asp Ser  
 1265 1270 1275 1280  
 Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys Arg Glu Asn  
 1285 1290 1295  
 Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu Lys Ala Gly Gln  
 1300 1305 1310  
 Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly Ala Pro Ala Ala His  
 1315 1320 1325  
 Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu Cys Pro Gly Pro Gln Leu  
 1330 1335 1340  
 Asn Thr Gly Thr Gln Leu Val Leu Gln His Val Gln Ala Leu Leu Val  
 1345 1350 1355 1360  
 Lys Arg Phe Gln His Thr Ile Arg Ser His Lys Asp Phe Leu Ala Gln

| 1365                                                                    | 1370 | 1375 |
|-------------------------------------------------------------------------|------|------|
| Ile Val Leu Pro Ala Thr Phe Val Phe Leu Ala Leu Met Leu Ser Ile<br>1380 | 1385 | 1390 |
| Val Ile Leu Pro Phe Gly Glu Tyr Pro Ala Leu Thr Leu His Pro Trp<br>1395 | 1400 | 1405 |
| Ile Tyr Gly Gln Gln Tyr Thr Phe Phe Ser Met Asp Glu Pro Gly Ser<br>1410 | 1415 | 1420 |
| Glu Gln Phe Thr Val Leu Ala Asp Val Leu Leu Asn Lys Pro Gly Phe<br>1425 | 1430 | 1435 |
| Gly Asn Arg Cys Leu Lys Glu Gly Trp Leu Pro Glu Tyr Pro Cys Gly<br>1445 | 1450 | 1455 |
| Asn Ser Thr Pro Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr Gln<br>1460 | 1465 | 1470 |
| Leu Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro Ser Pro Ser Cys<br>1475 | 1480 | 1485 |
| Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro Glu<br>1490 | 1495 | 1500 |
| Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser Thr Glu<br>1505 | 1510 | 1515 |
| Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile Ser Asp Phe Leu Val Lys<br>1525 | 1530 | 1535 |
| Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser Lys Phe Trp Val<br>1540 | 1545 | 1550 |
| Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly Gly Lys Leu Pro Val<br>1555 | 1560 | 1565 |
| Val Pro Ile Thr Gly Glu Ala Leu Val Gly Phe Leu Ser Asp Leu Gly<br>1570 | 1575 | 1580 |
| Arg Ile Met Asn Val Ser Gly Gly Pro Ile Thr Arg Glu Ala Ser Lys<br>1585 | 1590 | 1595 |
| Glu Ile Pro Asp Phe Leu Lys His Leu Glu Thr Glu Asp Asn Ile Lys<br>1605 | 1610 | 1615 |
| Val Trp Phe Asn Asn Lys Gly Trp His Ala Leu Val Ser Phe Leu Asn<br>1620 | 1625 | 1630 |
| Val Ala His Asn Ala Ile Leu Arg Ala Ser Leu Pro Lys Asp Arg Ser<br>1635 | 1640 | 1645 |
| Pro Glu Glu Tyr Gly Ile Thr Val Ile Ser Gln Pro Leu Asn Leu Thr<br>1650 | 1655 | 1660 |
| Lys Glu Gln Leu Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp Ala<br>1665 | 1670 | 1675 |
|                                                                         |      | 1680 |

Val Val Ala Ile Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala Ser  
 1685 1690 1695  
 Phe Val Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His Leu  
 1700 1705 1710  
 Gln Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn Phe  
 1715 1720 1725  
 Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val Gly  
 1730 1735 1740  
 Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu Asn Leu  
 1745 1750 1755 1760  
 Pro Ala Leu Val Ala Leu Leu Leu Leu Tyr Gly Trp Ala Val Ile Pro  
 1765 1770 1775  
 Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro Ser Thr Ala Tyr  
 1780 1785 1790  
 Val Ala Leu Ser Cys Ala Asn Leu Phe Ile Gly Ile Asn Ser Ser Ala  
 1795 1800 1805  
 Ile Thr Phe Ile Leu Glu Leu Phe Asp Asn Asn Arg Thr Leu Leu Arg  
 1810 1815 1820  
 Phe Asn Ala Val Leu Arg Lys Leu Leu Ile Val Phe Pro His Phe Cys  
 1825 1830 1835 1840  
 Leu Gly Arg Gly Leu Ile Asp Leu Ala Leu Ser Gln Ala Val Thr Asp  
 1845 1850 1855  
 Val Tyr Ala Arg Phe Gly Glu Glu His Ser Ala Asn Pro Phe His Trp  
 1860 1865 1870  
 Asp Leu Ile Gly Lys Asn Leu Phe Ala Met Val Val Glu Gly Val Val  
 1875 1880 1885  
 Tyr Phe Leu Leu Thr Leu Leu Val Gln Arg His Phe Phe Leu Ser Gln  
 1890 1895 1900  
 Trp Ile Ala Glu Pro Thr Lys Glu Pro Ile Val Asp Glu Asp Asp Asp  
 1905 1910 1915 1920  
 Val Ala Glu Glu Arg Gln Arg Ile Ile Thr Gly Gly Asn Lys Thr Asp  
 1925 1930 1935  
 Ile Leu Arg Leu His Glu Leu Thr Lys Ile Tyr Leu Gly Thr Ser Ser  
 1940 1945 1950  
 Pro Ala Val Asp Arg Leu Cys Val Gly Val Arg Pro Gly Glu Cys Phe  
 1955 1960 1965  
 Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Thr Thr Thr Phe Lys Met  
 1970 1975 1980

Leu Thr Gly Asp Thr Thr Val Thr Ser Gly Asp Ala Thr Val Ala Gly  
 1985 1990 1995 2000

Lys Ser Ile Leu Thr Asn Ile Ser Glu Val His Gln Asn Met Gly Tyr  
 2005 2010 2015

Cys Pro Gln Phe Asp Ala Ile Asp Glu Leu Leu Thr Gly Arg Glu His  
 2020 2025 2030

Leu Tyr Leu Tyr Ala Arg Leu Arg Gly Val Pro Ala Glu Glu Ile Glu  
 2035 2040 2045

Lys Val Ala Asn Trp Ser Ile Lys Ser Leu Gly Leu Thr Val Tyr Ala  
 2050 2055 2060

Asp Cys Leu Ala Gly Thr Tyr Ser Gly Gly Asn Lys Arg Lys Leu Ser  
 2065 2070 2075 2080

Thr Ala Ile Ala Leu Ile Gly Cys Pro Pro Leu Val Leu Leu Asp Glu  
 2085 2090 2095

Pro Thr Thr Gly Met Asp Pro Gln Ala Arg Arg Met Leu Trp Asn Val  
 2100 2105 2110

Ile Val Ser Ile Ile Arg Glu Gly Arg Ala Val Val Leu Thr Ser His  
 2115 2120 2125

Ser Met Glu Glu Cys Glu Ala Leu Cys Thr Arg Leu Ala Ile Met Val  
 2130 2135 2140

Lys Gly Ala Phe Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser Lys  
 2145 2150 2155 2160

Phe Gly Asp Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys Asp  
 2165 2170 2175

Asp Leu Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly Asn  
 2180 2185 2190

Phe Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln Phe  
 2195 2200 2205

Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu Ser  
 2210 2215 2220

His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln Thr Thr  
 2225 2230 2235 2240

Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr Glu Ser His  
 2245 2250 2255

Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser Arg Gln Ala Gln  
 2260 2265 2270

Asp

<210> 13

<211> 11  
<212> PRT  
<213> Human immunodeficiency virus type 1

<400> 13  
Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg  
1 5 10

<210> 14  
<211> 15  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: internalizing  
domain deruved from HIV tat protein

<400> 14  
Gly Gly Gly Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg  
1 5 10 15

<210> 15  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR primer  
2476-84

<400> 15  
tccatcttga gtggcctctt ccc 23

<210> 16  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR primer  
2476-85

<400> 16  
cagacccttc agccgcccac ag 22

<210> 17  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR primer  
2521-12

<400> 17  
atggccttct ggacacagct gatg 24

<210> 18  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR primer  
2508-66

<400> 18  
cttcaggcgt ctccagagca gg 22

<210> 19  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR primer  
2508-63

<400> 19  
gtgccctggc tccagggtct c 21

<210> 20  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR primer  
2477-29

<400> 20  
cacgtagggc aggtagagg 19

<210> 21  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR primer  
2520-24

<400> 21  
gaaccgccca ttcaccatga tgg 23

<210> 22  
<211> 27

<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR primer  
2533-69

<400> 22  
atggctttct gcacacagtt gatgctc 27

<210> 23  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR primer  
2507-27

<400> 23  
ccaggccaca cacagtacgt agg 23

<210> 24  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR primer  
2507-28

<400> 24  
acacagtacg tagggcagat agag 24